



HAL
open science

Caractérisation et modification de lectines fongiques pour la reconnaissance spécifique de motifs glucidiques principalement associés aux cancers

Aurore Cabanettes

► **To cite this version:**

Aurore Cabanettes. Caractérisation et modification de lectines fongiques pour la reconnaissance spécifique de motifs glucidiques principalement associés aux cancers. Biochimie [q-bio.BM]. Université Grenoble Alpes, 2019. Français. NNT : 2019GREAV021 . tel-03381202

HAL Id: tel-03381202

<https://theses.hal.science/tel-03381202>

Submitted on 16 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

Pour obtenir le grade de

DOCTEUR DE LA COMMUNAUTÉ UNIVERSITÉ GRENOBLE ALPES

Spécialité : Chimie Biologie

Arrêté ministériel : 25 mai 2016

Présentée par

Aurore CABANETTES

Thèse dirigée par **Annabelle VARROT**

Préparée au sein du **Laboratoire Centre de Recherche sur les
MACromolécules Végétales**
dans l'**École Doctorale Chimie et Sciences du Vivant**

**Caractérisation et modification de lectines
fongiques pour la reconnaissance spécifique
de motifs glucidiques principalement
associés aux cancers**

**Characterization and modification of fungal
lectins for specific binding to carbohydrates
patterns mainly involved in cancers**

Thèse soutenue publiquement le **15 Octobre 2019**,
devant le jury composé de :

Madame ANNABELLE VARROT

CHARGE DE RECHERCHE, CNRS DELEGATION ALPES, Directeur de
thèse

Madame JULIE BOUCKAERT

CHARGE DE RECHERCHE, CNRS DELEGATION HAUTS-DE-FRANCE,
Rapporteur

Monsieur JEROME NIGOU

DIRECTEUR DE RECHERCHE, CNRS DELEGATION OCCITANIE
OUEST, Rapporteur

Madame MURIEL BARDOR

PROFESSEUR, UNIVERSITE DE ROUEN-NORMANDIE, Examineur

Monsieur ROMAIN VIVES

DIRECTEUR DE RECHERCHE, CNRS DELEGATION ALPES,
Président



REMERCIEMENTS

Arrivée au CERMAV il y a maintenant 4 ans en 2015 pour un premier stage de Master, j'ai eu la chance de rencontrer un grand nombre de personnes que je tiens à remercier.

Je remercie tout d'abord l'Université Grenoble Alpes pour avoir financé mes trois années de thèse.

Je remercie le Dr. Julie Bouckaert et le Dr. Jérôme Nigou pour avoir accepté d'évaluer ce travail et pour la qualité et la pertinence de leurs remarques. Je remercie également le Dr. Muriel Bardor et le Dr. Romain Vives d'avoir accepté d'être membres de ce jury, et de l'intérêt qu'ils ont manifesté pour mon travail.

Je souhaite adresser des remerciements particuliers à ma directrice de thèse, le Dr. Annabelle Varrot, grâce à qui j'ai pu intégrer le CERMAV en 2015 pour mon premier stage de Master 1. Je la remercie pour tout ce qu'elle m'a apporté pendant toutes ces années, pour sa bienveillance, ses conseils, son implication. Je la remercie de m'avoir toujours soutenue et en particulier lors de la préparation au concours de l'école doctorale qui m'a permis de réaliser un travail de thèse sur ce projet qui me tenait à cœur. Je la remercie pour tous les moments partagés et ses grandes qualités humaines.

Je remercie également toutes les personnes avec qui nous avons collaboré dans le cadre de ce travail. Merci au Dr. Carlo Unverzagt, Carolina Spies et Lukas Perkams à l'université de Bayreuth en Allemagne. Merci au Dr. Jacques Lependu et au Dr. Adrien Breiman à l'université de Nantes pour leur implication dans ce projet. Merci au Dr. Jean-Luc Coll et à Maxime Henry de m'avoir si bien accueillie au sein de l'IAB et de m'avoir permis d'élargir mes connaissances scientifiques dans un domaine qui n'était pas le mien. Merci également au Dr. Jérica Sabotic et au Dr. Anja Klančnik à Ljubljana pour leur accueil chaleureux au sein de leur université et au sein de leur beau pays.

Je voudrais aussi remercier tous les membres de l'équipe Glycobiologie Moléculaire et Structurale pour tout ce qu'ils m'ont apporté :

Un grand merci à Valérie Chazalet et Emilie Gillon pour leur aide précieuse au laboratoire et leur bonne humeur au quotidien.

Merci au Dr. Olivier Lerouxel pour son énergie positive, ses bons conseils et sa sympathie. Merci également au Dr. Anne Imberty, au Dr Serge Perez et au Dr. Christelle Breton qui nourrissent l'équipe de la richesse de leurs connaissances scientifiques.

Merci à tous autres membres de l'équipe GBMS, à ceux qui n'ont été que de passage comme à ceux qui viennent de rejoindre le groupe. Merci de contribuer à cette belle ambiance qui règne et perdure dans cette équipe quelles que soient les personnes qui la composent et l'enrichissent au fil des années.

Je remercie bien sûr Oriane Machon et François Portier, qui au-delà de relations professionnelles remarquables sont devenus des amis sincères. Merci Oriane pour tous les bons moments qu'on a partagé dans le même bureau, à nos longues discussions sur la science comme sur la vie. Merci François pour la personne que tu es, ton ouverture d'esprit à tout point de vue et ta sensibilité. Merci pour la richesse de nos discussions et cette belle amitié qui est née.

Merci également aux autres membres du laboratoire avec qui j'ai eu la chance d'échanger au cours de ces années. Je ne pourrai pas citer tous les noms mais je retiendrai beaucoup de bienveillance et de sympathie de la part de toutes ces personnes qui se reconnaîtront.

Enfin, je remercie mes amis et ma famille pour leur amour et leur soutien quotidien et sans faille.

Merci Meryem pour tout.

Merci à ma grand-mère Mireille.

SOMMAIRE

Liste des abréviations.....	12
Liste des figures.....	14
Liste des tableaux.....	20
I. INTRODUCTION	21
1. Glycobiologie.....	23
1.1 Généralités : glycobiologie et glycosylation	23
1.2 La glycosylation des protéines	24
1.2.1 La N-glycosylation des eucaryotes.....	25
1.2.2 La O-glycosylation	27
1.2.3 Les glycosyltransférases	28
1.3 Les anomalies de la glycosylation	30
1.3.1 Les maladies héréditaires liées aux défauts de glycosylation	30
1.3.2 Altérations de la glycosylation dans les cancers	31
1.3.2.1 Généralités	31
1.3.2.2 La core fucosylation.....	33
1.3.2.3 L'antigène TF (Thomsen–Friedenreich)	35
1.3.3 Glycosylation et efficacité de certains anticorps thérapeutiques.....	36
2. Les lectines : outils performants d'analyse de la glycosylation	38
2.1 Généralités sur les lectines	38
2.2 Les lectines fongiques.....	40
2.3 Utilisation des lectines comme outil de reconnaissance des sucres	42
2.3.1 Analyse de la glycosylation.....	42
2.3.2 Diagnostic de cancers	45
2.4 Les lectines recombinantes.....	47
2.4.1 Intérêt et mise en pratique	47
2.4.2 Particularité des lectines à pont disulfure	48
3. Les lectines spécifiques de marqueurs cancéreux.....	50
3.1 Reconnaissance de la core-fucosylation	50
3.1.1 État de l'art	50
3.1.2 PhoSL et PhoNL.....	54
3.2 Reconnaissance de l'antigène TF	55
3.2.1 État de l'art	55
3.2.2 TrfbL2	57
II. OBJECTIFS DE LA THÈSE	59

III. MATÉRIEL ET MÉTHODES.....	63
1. Clonage et mutagenèse dirigée	65
2. Production et purification.....	68
2.1 Production et extraction	68
2.1.1 rPhoSL, rPhoNL, rPhoSLm, rPhoNLm	68
2.1.2 rPhoNLmSS	69
2.1.3 rTrfbL2 et ses mutants	69
2.1.4 rPhoNLSeMet.....	70
2.2 Purification	70
2.2.1 rPhoSL, rPhoNL, et leurs mutants.....	70
2.2.2 rTrfbL2-WT, rTrfbL2-Y111Q, rTrfbL2-Y111A, rTrfbL2-R100G	71
2.2.3 rTrfbL2-Y111QD77A, rTrfbL2-D77A.....	71
3. Caractérisation biochimique	72
3.1 Électrophorèse sur gel de polyacrylamide	72
3.1.1 Gel SDS-PAGE Tris-glycine.....	72
3.1.2 Gel SDS-PAGE Tris-tricine	73
3.2 Chromatographie d'exclusion de taille	74
3.3 SEC-MALS	75
3.4 Diffusion dynamique de la lumière (DLS).....	76
3.5 Test de stabilité thermique	76
4. Marquage des lectines.....	77
4.1 Alexa 488, Alexa 680 et biotine	77
4.2 Peroxydase.....	77
5. Étude de l'interaction lectines-sucre.....	77
5.1 Tests d'hémagglutination de globules rouges de lapin	77
5.2 Analyse de la spécificité sur puces à sucres (glycan array)	78
5.3 Tests d'interaction ELLA (Enzyme-Linked Lectin Assay)	78
5.4 Titration calorimétrique isotherme (ITC).....	79
5.5 Résonance plasmonique de surface (SPR).....	81
5.6 Bio-Layer Interferometry (BLI)	82
6. Étude structurale	84
6.1 N-glycanes synthétisés pour l'étude	84
6.2 Principes de la cristallographie aux rayons X	84
6.2.1 La cristallogénèse	84
6.2.2 Phénomène de diffraction des rayons X	86

6.2.3	Le problème des phases	88
6.2.4	L'affinement.....	89
6.3	Cristallisation et collecte des données	90
6.4	Résolution des structures et affinement	91
7.	Mesure de l'activité anti-adhésive pour des micro-organismes.....	91
8.	Étude biologique sur cellules et tissus cancéreux.....	92
8.1	Cytométrie en flux	92
8.1.1	Principe de la cytométrie en flux.....	92
8.1.2	Traitement des échantillons.....	93
8.2	Lectine-histochimie	95
IV.	RÉSULTATS ET DISCUSSION	97
1.	Outil de reconnaissance de l'antigène TF : rTrfbL2	100
1.1	Caractéristiques de TrfbL2 et production sous forme recombinante.....	100
1.2	Conception de mutants abrogeant la reconnaissance du GlcNAc	104
1.2.1	Production et purification de rTrfbL2-Y111A et rTrfbL2-Y111Q	105
1.2.2	Étude structurale du mutant Y111A	105
1.2.3	Essais de purification de rTrfbL2-Y111QD77A et rTrfbL2- D77A	107
1.2.4	Caractérisation du mutant R100G : stabilité et activité	112
2.	Reconnaissance de la core-fucosylation : rPhoSL.....	118
2.1	Production et purification.....	118
2.2	Analyse de l'état d'oligomérisation, de la stabilité, et de l'activité.....	120
2.3	Étude de l'interaction lectine-sucre	121
2.3.1	Analyse de la spécificité sur puces à sucres (glycan array)	121
2.3.2	Mesure de l'affinité par Titration Calorimétrique Isotherme (ITC).....	122
2.3.3	Tests en SPR (Résonance Plasmonique de Surface) et BLI (Bio-Layer Interferometry).....	124
2.3.4	Tests d'interaction avec des glycoprotéines en ELLA (Enzyme-linked lectin assay).....	125
2.4	Étude des propriétés anti-adhésives de rPhoSL sur des micro-organismes	126
2.5	Étude structurale	127
2.5.1	Cristallogénèse	127
2.5.2	Résolution de la structure de rPhoSL en absence de ligand.....	128
2.5.3	Résolution de la structure de rPhoSL en complexe avec des N-glycanes core-fucosylés.....	132
2.6	Définition d'une nouvelle famille de lectines	139
2.7	Mutations de rPhoSL (rPhoSLm), purification et analyse.....	140

3. Reconnaissance de la core-fucosylation : rPhoNL	144
3.1 Production et purification de rPhoNL.....	144
3.2 Analyse de l'état d'oligomérisation, de la stabilité et de l'activité	145
3.3 Étude de l'interaction lectine-sucre	147
3.3.1 Analyse de la spécificité sur puces à sucres (glycan array)	147
3.3.2 Mesure de l'affinité par Titration Calorimétrique Isotherme.....	148
3.1 Étude structurale	148
3.1.1 Description de la structure de rPhoNL.....	149
3.1.2 Description des interactions avec le N-glycane 1.....	151
3.2 Mutations pour améliorer l'affinité de rPhoNL pour ses ligands (rPhoNLm).....	156
3.2.1 Production et analyse biochimique de rPhoNLm.....	157
3.2.2 Analyse structurale de rPhoNLm avec le N-glycane 6.....	158
3.2.3 Analyse de l'état d'oligomérisation de rPhoNLm.....	160
3.2.4 Conclusions	163
3.3 Mutations pour créer un dimère stable de rPhoNL : rPhoNLmSS	164
3.4 Étude comparative des protéines strictement spécifiques de la core-fucosylation.....	166
4. Étude biologique sur cellules et tissus	171
4.1 Cytométrie en flux sur lignées cancéreuses	171
4.1.1 PhoSL, PhoSLm et PhoNLm	171
4.1.2 rTrfbL2 et rTrfbL2-Y111A.....	175
4.1.3 Conclusion.....	177
4.2 Cytométrie en flux sur cellules mononucléées du sang périphérique.....	177
4.3 Lectine-histochimie	178
4.3.1 Tests sur tissus humains	179
4.3.2 Tests sur tumeurs induites chez le rat	181
4.3.2.1 rPhoSLm et rPhoNLm.....	182
4.3.2.2 rTrfbL2 t rTrfbL2-Y111A.....	184
4.3.3 Tests d'inhibition.....	185
4.4 Conclusions générales de l'étude biologique sur cellules et tissus	187
V. CONCLUSIONS ET PERSPECTIVES	189
VI. BIBLIOGRAPHIE	195
VII. ANNEXES	213
1. Récapitulatif des séquences des oligonucléotides utilisés lors des expériences de mutagenèse dirigée.....	215
2. Description détaillée des structures des N-glycanes utilisés dans cette étude	216

3. Représentation des N-glycanes 1, 2, 3, 5 et 6 en densité électronique	217
4. Analyse des angles de torsion des N-glycanes	221

LISTE DES ABRÉVIATIONS

AAL	<i>Aleuria auratia</i> lectin
ABL	<i>Agaricus bisporus</i> lectin
AFP	Alpha-fétoprotéine
AOL	<i>Aspergillus oryzae</i> lectin
ARN	Acide RiboNucléique
Å	Ångström
APS	Ammonium persulfate
BEL	<i>Boletus edulis</i> lectin
BLI	Bio-Layer Interferometry
BSA	Bovine Serum Albumin
BTL	<i>Bryothamnion triquetrum</i> lectin
CAZy	Carbohydrate Active Enzyme
CDG	Congenital Disorders of Glycosylation
CSL	<i>Cephalosporium curvulum</i> lectin
DLS	Dynamic Light Scattering
DO	Densité Optique
DTT	Dithiothréitol
EDC	1-éthyl-3-(3-diméthylaminopropyl) carbodiimide
EDTA	Ethylenediaminetetraacetic Acid
ELLA	Enzyme-linked lectin assay
FACS	Fluorescence Activated Cell Sorting
Fuc	Fucose
FUT	Fucosyltransférase
GT	Glycosyl Transférases
Gal	Galactose
GalNAc	N-Acétylgalactosamine
GlcNAc	N-Acétylglucosamine
GSH	Glutathion réduit
GSSG	Glutathion oxydé
HCC	Carcinome hépatocellulaire
Hpt	Haptoglobine
IPTG	IsoPropyl-b-D-ThioGalactopyranoside
ITC	Tritration Calorimétrique Isotherme
K_A	Constante d'association à l'équilibre
k_a	Constante de vitesse d'association
K_D	Constante de dissociation à l'équilibre
k_d	Constante de vitesse de dissociation
kDa	kilodalton
LB	Lysogen Broth
LCA	<i>Lens culinaris</i> agglutinin

mA	milliampère
MAD	Multi wavelength Anomalous Scattering
MALS	Multiangle Light Scattering
Man	Mannose
MES	acide 2-(N-morpholino)éthanosulfonique
MIR	Multiple Isomorphous Replacement
<i>n</i>	stœchiométrie
NeuAc	Acide N-acétyl-Neuraminique
NHS	N-hydroxysuccinimide
OPD	o-phenylenediamine dihydrochloride
PBS	Phosphate Buffer Saline
PBS-T	Phosphate Buffer Saline-Tween
PCR	Réaction de Polymérisation en Chaîne
PDB	Protein Data Bank
PhoNL	<i>Pholiota nameko</i> lectin
PhoSL	<i>Pholiota squarrosa</i> lectin
PSA	Prostate-specific antigen
PSA	<i>Pisum sativum</i> agglutinin
PVL	<i>Psathyrella velutina</i>
RMN	Résonance magnétique nucléaire
RSL	<i>Rhizopus stolonifer</i> lectin
qRT-PCR	real-time reverse transcription-PCR
RU	Unité de résonance
SAD	Single wavelength Anomalous Scattering
SDS	Sodium dodecyl Sulfate
SDS-PAGE	Sodium dodecyl Sulfate PolyAcrylamide Gel Electrophoresis
SIR	Single Isomorphous Replacement
siRNA	petit ARN interférent
SL2-1	<i>Streptomyces rapamycinicus</i> lectin
SNFG	Symbols Nomenclature For Glycans
SPR	Résonance plasmonique de surface
SRL	<i>Sclerotium rofsii</i> lectin
TACA	Tumor Associated Carbohydrate Antigen
TEMED	Tetramethylethylenediamine
TEV	Tobacco Etch Virus
T_m	Melting temperature
TrfbL2	<i>Terfezia boudieri</i> lectin 2
Tris	trishydroxyméthylaminométhane
TSA	Thermal Shift Assay
TF	Thomsen-Friedenreich
ΔG	Variation d'énergie libre
ΔH	Enthalpie
ΔS	Entropie
XCL	<i>Xercomus chrysenteron</i> lectin

LISTE DES FIGURES

Figure I-1. A. Description des anomères alpha et beta du D-glucose. B. Représentation symbolique de quelques monosaccharides (nomenclature SNFG).....	23
Figure I-2. Représentation de la surface cellulaire et des glycoconjugués présents (A) et description des rôles des sucres dans les cellules (B).	24
Figure I-3. Représentation schématique du noyau des N-glycanes et des trois types de N-glycanes chez l’homme.	26
Figure I-4. Schéma simplifié des grandes étapes de la N-glycosylation selon les espèces.	27
Figure I-5. Description des huit noyaux de la O-glycosylation de type mucine et de quelques exemples d’élancements possibles.	28
Figure I-6. Réactions des glycosyltransférases impliquant l’inversion ou la rétention du centre anomérique du sucre donneur.	29
Figure I-7. Structures de glycosyltransférases de types A et B.	29
Figure I-8. Dissociation des cellules tumorales et invasion conduisant à la formation de tumeurs métastatiques d’après (Pinho and Reis, 2015).....	31
Figure I-9. Représentation de motifs de N- et O- glycosylation altérés retrouvés dans les cancers.	32
Figure I-10. Voies de synthèse du GDP fucose et réaction de core-fucosylation en α 1,6 des N-glycanes par la fucosyltransférase 8 (Fut8) dans l’appareil de Golgi.	34
Figure I-11. A. Représentation schématique d’une immunoglobuline (Saxena and Wu, 2016). B. Schéma des principaux mécanismes d’action des anticorps.....	37
Figure I-12. Principaux rôles des lectines <i>in vivo</i>	39
Figure I-13. Récapitulatif des divers rôles biologiques des lectines fongiques et de leurs applications en biologie et en recherche biomédicale.	41
Figure I-14. Approches directes (A) et hybrides (B) des tests ELLA (Enzyme Linked Lectin Assay). D’après Hashim et al., 2017.....	44
Figure I-15. Représentation schématique d’une expérience sur micro-puce à lectine.	44
Figure I-16. Illustration des différents biocapteurs à base de lectine.	45
Figure I-17. Représentation schématique de transferts d’électrons entre une protéine substrat (ronds gris) et une oxydoréductase (lignes noires) permettant l’oxydation, la réduction ou l’isomérisation des ponts disulfures. D’après (Berkmen, 2012).....	49
Figure I-18. Réactions d’oxydation et d’isomérisation des ponts disulfures dans le périplasme d’ <i>E. coli</i> . D’après de Marco, 2009.....	50

Figure I-19. Structure de AAL (<i>Aleuria auratia</i>) et AOL (<i>Aspergillus oryzae</i>).....	51
Figure I-20. Superposition des structures de LCA et PSA.	52
Figure I-21. Schématisation des différences de structure primaire et tridimensionnelle entre PhoNL et PhoSL (A) et alignement de séquence avec RSL et SL2-1 (B).	55
Figure I-22. Structure du tétramère de ABL avec un zoom sur un des monomères.	56
Figure I-23. Structure du tétramère de TrfbL2 (A) et d'un monomère avec les sucres en interaction sur les sites à GlcNAc et à GalNAc (B).	58
Figure III-1. A. Carte du vecteur pET39b (Invitrogen) ; B. caractéristiques du vecteur modifié pET39b-TEV.	65
Figure III-2. Carte du vecteur pET29a.	66
Figure III-3. Plasmides utilisés pour l'expression des mutants de rTrfbL2. Clonage entre les sites NcoI et XhoI du site de clonage multiple.	67
Figure III-4. Principe du test ELLA.	79
Figure III-5. Schéma du fonctionnement d'un appareil d'ITC (A) et d'un thermogramme classique (B).	80
Figure III-6. Principe de la résonance plasmonique de surface (SPR).	81
Figure III-7. Principe du BLI.	83
Figure III-8. Ensemble des N-glycanes fucosylés en α -(1,6) utilisés dans l'étude.....	84
Figure III-9. Cristallogénèse : technique de diffusion de vapeur en goutte suspendue.....	85
Figure III-10. Diagramme de phase et représentation des conditions nécessaires à la cristallisation par la technique de diffusion de vapeur.....	85
Figure III-11. Diffraction des rayons X par un cristal à l'origine du cliché de diffraction.	86
Figure III-12. Principe du FACS.....	93
Figure IV-1. Alignement des séquences homologues à TrfbL2 de structure connue.	101
Figure IV-2. A. Structure du tétramère de TrfbL2 native. B. Description de la structure d'un monomère de TrfbL2 native.	101
Figure IV-3. Description des interactions entre TrfbL2-WT et ses ligands.	102
Figure IV-4. Structure du GlcNAc (A) et du GalNAc (B).	103
Figure IV-5. A. Interaction de rTrfbL2-WT avec du GlcNAc β 1-3Gal. B. Schéma d'un N-glycane possédant le disaccharide GlcNAc β 1-3Gal.	104
Figure IV-6. Interaction du GlcNAc avec rTrfbL2-Y111A (monomère).....	107

Figure IV-7. Thermal Shift Assay avec rTrfbL2-WT et une gamme de tampons de pH 5 à 10.	108
Figure IV-8. Gel SDS-PAGE Tris glycine 15% de la purification de rTrfbL2-Y111QD77A cloné dans le vecteur pET32-TEV sur Histrap.	110
Figure IV-9. Gel SDS-PAGE Tris glycine 15% de la deuxième étape de purification de rTrfbL2-Y111QD77A cloné dans le vecteur pET32-TEV sur Histrap (après coupure de l'étiquette Nter).	111
Figure IV-10. Description des interactions de la chaîne latérale de l'aspartate 77.....	112
Figure IV-11. A. Gel SDS-PAGE tris-glycine 15% de la purification de rTrfbL2-R100G sur colonne d'affinité au GlcNAc.....	113
Figure IV-12. Test de stabilité avec rTrfbL2-WT, rTrfbL2-Y111A et rTrfbL2-R100G.	114
Figure IV-13. Comparaison des capacités d'héماغglutination de rTrfbL2-WT et de rTrfbL2- R100G.	114
Figure IV-14. Description des interactions de la chaîne latérale de l'arginine 100.	115
Figure IV-15. Gels SDS-PAGE des deux étapes de purification de rPhoSL sur IMAC et DEAE.	119
Figure IV-16. Ajout des trois acides aminés « GAM » lors des étapes de clonage dans le vecteur Pet39-TEV.....	119
Figure IV-17. Test d'héماغglutination de rPhoSL sur globules rouges de lapin avec du sang préalablement traité (T) ou non (NT) à la trypsine.	120
Figure IV-18. Test TSA avec rPhoSL et 26 tampons de pH 5 à PH 10.	121
Figure IV-19. Résultats de l'étude de rPhoSL sur puce à sucres version 5.2.....	122
Figure IV-20. Résultat des mesures d'affinité de rPhoSL pour le N-glycane 1 (A) et le chitobiose core-fucosylé (B) par ITC.....	123
Figure IV-21. Alignement de la séquence de rPhoSL avec celles de PhoSL native et des trois répétitions de SL2-1.	123
Figure IV-22. Fixation de rPhoSL biotinylée su des glycoprotéines par ELLA.....	126
Figure IV-23. Photos des cristaux de rPhoSL (A) et de rPhoSL-ligands (B).	127
Figure IV-24. Représentation schématique de la structure rPhoSL native (β -prisme-III).	129
Figure IV-25. Description détaillée de la structure de rPhoSL.....	130
Figure IV-26. Différences structurales entre rPhoSL (β -prisme III) et GNA (β -prisme II)...	131
Figure IV- 27. Superposition des structures de rPhoSL et des structures RMN et cristallographiques du peptide synthétique.	132

Figure IV-28. Schéma des N-glycanes avec lesquels rPhoSL a été co-cristallisée et dont la structure a été résolue.....	134
Figure IV-29. A. Superposition des trimères de rPhoSL en absence de ligand avec les structures de rPhoSL en interaction avec les N-glycanes 1, 2, 3 et 5. B. Représentation du N-glycane 1 à chaque interface entre deux monomères.....	134
Figure IV-30. Description des interactions entre rPhoSL et le N-glycane 3.	135
Figure IV-31. A. Description des interactions avec le N-glycane 1. B. Comparaison des interactions entre le N-glycane 1 (bleu) et le N-glycane 3 (jaune).	136
Figure IV-32. Flexibilité des N-glycanes.	137
Figure IV-33. Description des interactions avec le N-glycane 5 en densité. Coloration des acides aminés par monomère. Représentation schématique à droite.	138
Figure IV-34. Alignement multiple de rPhoSL.....	139
Figure IV-35. Comparaison des capacités d'hémagglutination des globules rouges de lapin entre rPhoSLm et rPhoSL.....	141
Figure IV-36. Alignement des séquences de rPhoSL et de rPhoNL.	144
Figure IV-37. Gels SDS-PAGE des deux étapes de purification de rPhoNL sur IMAC et DEAE.	145
Figure IV-38. Test d'hémagglutination sur globules rouges de lapin de rPhoNL pour du sang préalablement traité (T) ou non (NT) à la trypsine. Comparaison avec rPhoSL.....	146
Figure IV-39. Tests TSA avec rPhoNL et 26 tampons de pH 5 à pH 10.	147
Figure IV-40. Résultats de l'étude de rPhoNL sur puce à sucre version 5.2.....	148
Figure IV-41. Photo de cristaux de rPhoNL co-cristallisée avec le N-glycane 1.	149
Figure IV-42. Représentation de la structure de rPhoNLSeMet résolue en co-cristallisation avec le glycane 1 (β -prisme III).....	151
Figure IV-43. Analyse des interactions entre rPhoNLSeMet et le N-glycane 1.....	152
Figure IV-44. Représentation de l'encombrement stérique empêchant la fixation des sucres au niveau des boucles a et b de rPhoNLSeMet.	154
Figure IV-45. Alignement des séquences de rPhoNL et SL2-1 par répétition (1, 2, 3 pour rPhoNL et a, b, c pour SL2-1).	155
Figure IV-46. Alignement de séquences centré sur les grandes boucles de surface.....	156
Figure IV-47. Alignement des séquences de rPhoNL et rPhoSL.....	157
Figure IV-48. Superposition des structures de rPhoNL et rPhoNLm (A) et descriptions des interactions de la chaîne A de rPhoNLm avec le N-glycane 6 (B).....	160
Figure IV-49. Représentation de l'interface entre les monomères de rPhoNLm.....	161

Figure IV-50. Analyse par filtration sur gel de rPhoNLm sur colonne Enrich 70.....	162
Figure IV-51. Analyse par filtration sur gel de rPhoNLm sur colonne S200.....	162
Figure IV-52. Expérience de SEC-MALS avec rPhoNLm et une colonne S200 increase.....	163
Figure IV-53. Analyse sur gel SDS-PAGE de la dimérisation de rPhoNLmSS purifiée.....	165
Figure IV-54. Alignement des séquences de rPhoSL et de PhoSLc.....	167
Figure IV-55. Superposition des structures de rPhoSL et de la structure de PhoSLc.....	167
Figure IV-56. Description des orientations différentes de l'alanine 1 entre rPhoSL (6fx3 chaîne A) et PhoSLc (6a87 chaîne D).....	168
Figure IV-57. Comparaison de la reconnaissance de différentes lignées cellulaires par rPhoSL, rPhoSLm et rPhoNLm.....	171
Figure IV-58. Inhibition de la fucosylation α 1,6 des N-glycanes sur les cellules A549 et MCF7 avec des siRNA FUT8.....	173
Figure IV-59. Inhibition de la fixation de rPhoSLm, rPhoSL et rPhoNLm sur cellules MINO, Z138 et HT-29 en utilisant le 6-AF ou du chitobiose core-fucosylé.....	174
Figure IV-60. Inhibition de la fixation de rPhoSLm et rPhoNLm sur cellules MCF7, A549 et HT-29 en utilisant 20 μ M de 6-AF.....	175
Figure IV-61. Étude de la fixation de rTrfbL2 et rTrfbL2-Y111A sur les cellules MCF7 en fonction de la concentration par la méthode FACS.....	176
Figure IV-62. Étude de la spécificité de reconnaissance de l'antigène TF sur les cellules MCF7 pour rTrfbL2 et rTrfbL2-Y111A.....	177
Figure IV-63. Expérience de cytométrie en flux de la reconnaissance de PBMCs par rPhoSL.	178
Figure IV-64. Marquage de coupes de tissus humains issus de cancers du côlon par rPhoSL, rPhoSLm et rPhoNLm.....	179
Figure IV-65. Tests d'inhibition du marquage de coupes tumorales coliques par rPhoNLm.	180
Figure IV-66. Marquage de tissus sains et comparaison de tissu sain et tumoraux coliques.	181
Figure IV-67. Comparaison du marquage de différents tissus par rPhoSLm et rPhoNLm....	183
Figure IV-68. Comparaison du marquage de différents tissus par rTrfbL2 et rTrfbL2-Y111A.	185
Figure IV-69. Tests d'inhibition de la reconnaissance des lectines rPhoSLm, rPhoNLm, rTrfbL2 et rTrfbL2-Y111A sur des coupes de tissus tumoraux.....	186

Figure VII-1. Densité électronique $2F_{\text{obs}}-F_{\text{calc}}$ du N-glycane 1 pour les 12 chaînes de rPhoSL affichée à 1σ	218
Figure VII-2. Densité électronique $2F_{\text{obs}}-F_{\text{calc}}$ des N-glycanes 2, 3 et 5 pour les chaînes A et B de rPhoSL affichée à 1σ	219
Figure VII-3. Densité électronique $2F_{\text{obs}}-F_{\text{calc}}$ du N-glycane 1 pour les chaînes A et B de rPhoNLSeMet et du N-glycane 6 pour les chaînes A et B de rPhonLm affichée à 1σ	220

LISTE DES TABLEAUX

Tableau I-1. Liste des familles dans lesquelles sont classées les lectines fongiques et des spécificités des lectines retrouvées dans chaque famille (Kobayashi and Kawagishi, 2014).....	42
Tableau III-1. Caractéristiques des vecteurs pProEx Htb, pET-TEV et pET32-TEV utilisés pour le clonage des mutants rTrfbL2-D77A et rTrfbL2-Y111QD77A.	67
Tableau III-2. Description des cellules utilisées dans cette étude.....	68
Tableau III-3. Ensemble des conditions testées pour la purification des mutants rTrfbL2-Y111QD77A et rTrfbL2-D77A.	72
Tableau III-4. Conditions pour la préparation de gels Tris-tricine 16.5 % et des tampons de migration 1X.	74
Tableau III-5. Description de la maille élémentaire, des sept systèmes cristallins, des 14 réseaux de bravais et quelques exemples de groupes d'espaces possibles pour chaque système. P : Primitif, I : corps centré, C : bases centrées, F : Faces centrées.	87
Tableau III-6. Milieux de cultures utilisés pour les différentes lignées cellulaires utilisées en cytométrie en flux	93
Tableau VII-1. Récapitulatif des amorces utilisées au cours de l'étude. Les codons mutés sont indiqués en rouge.	215
Tableau VII-2. Structures détaillées des N-glycanes utilisés dans l'étude.....	217
Tableau VII-3. Tableau des angles de torsion pour les N-glycanes en interaction avec rPhoSL (1, 2, 3 et 5).....	222

I.

INTRODUCTION

1. Glycobiologie

1.1 Généralités : glycobiologie et glycosylation

La glycobiologie a pour objectif d'étudier et de comprendre la structure, la synthèse, la dégradation et la modification des sucres ainsi que leurs fonctions biologiques. Les sucres peuvent être désignés par différents termes, tels que « carbohydrates », un mot issu du latin « carbo » signifiant « charbon » et du grec hydor signifiant « eau ». Cette appellation trouve son origine dans la nature des atomes qui composent les sucres : le carbone, l'oxygène et l'hydrogène, et est directement liée à la formule des sucres $C_n(H_2O)_m$ où $n \geq m$ (Gabiuis, 2018). Par ailleurs, on parle généralement de « glycane » lorsque les sucres sont liés à des protéines ou des lipides mais aussi de glucides (Taylor and Drickamer, 2003).

Les monomères de sucre, ou monosaccharides, sont la plus petite unité permettant ensuite de former de longues chaînes de glycanes que sont les oligosaccharides (moins de 12 monosaccharides) et les polysaccharides (plus de 12 monosaccharides) (Prestegard et al., 2015). Les monosaccharides sont reliés par des liaisons osidiques covalentes entre le carbone anomérique (C1 ou C2) du premier monosaccharide (sucre donneur) et n'importe quelle fonction hydroxyle non anomérique du monosaccharide suivant (sucre accepteur). Il existe à ce jour 150 monosaccharides identifiés qui peuvent être reliés entre eux par différents types de liaisons, former des structures branchées ou encore être modifiés (sulfation, phosphorylation, acétylation ou méthylation ...). Le carbone anomérique des sucres peut prendre deux positions : α ou β (**Figure I-1.A**). Les sucres existent sous forme libre ou liée à des protéines ou des lipides formant les glycoprotéines et les glycolipides appelés aussi glycoconjugués. Les sucres sont la classe de biomolécules la plus diverse et la plus complexe, ce qui a nécessité l'élaboration d'une représentation symbolique, appelée SNFG (Symbols Nomenclature For Glycans) dont les symboles des monosaccharides évoqués dans cette étude sont représentés dans la **Figure I-1.B** (Perez, 2018; Varki et al., 2015).

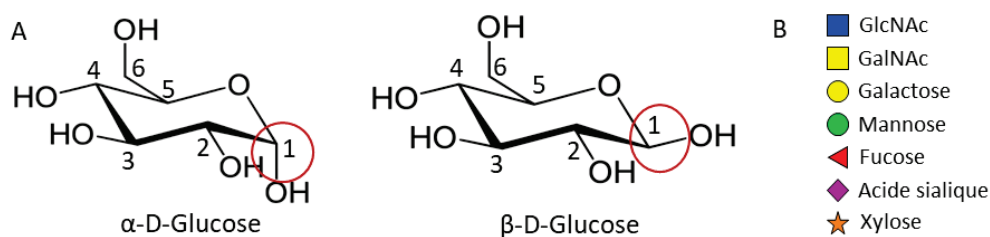


Figure I-1. A. Description des anomères alpha et beta du D-glucose. B. Représentation symbolique de quelques monosaccharides (nomenclature SNFG).

Les sucres sont des composants essentiels de tout organisme vivant et ont des rôles très variés qui vont bien au-delà de la production d'énergie résultant de la dégradation du glucose et du glycogène et des fonctions de structuration que tous les biochimistes connaissent. Ils ont des rôles dans l'adhésion cellulaire et la signalisation, dans le trafic de molécules dans les cellules et les organismes, dans la transduction de signaux ou encore dans les mécanismes d'endocytose et dans l'immunité innée. Certains sont d'ailleurs appelés antigènes lorsque qu'ils sont reconnus par les récepteurs de bactéries, virus, parasites ou toxines par exemple, ou encore lorsqu'ils sont reconnus par des anticorps comme les antigènes des groupes sanguins. Les sucres font partie intégrante de la matrice extracellulaire où ils jouent des rôles de barrière protectrice et d'organisation (**Figure I-2**) (Ohtsubo and Marth, 2006; Varki, 1993). Les glycoconjugués, en recouvrant la surface cellulaire, constituent la première couche d'informations reçues ou émises d'une cellule et le glycome forme le code barre des cellules. La grande diversité structurale des sucres permet d'accéder à une haute capacité de codage de l'information bien supérieure à celle observée pour les protéines et les acides nucléiques. Le langage des sucres ou glycode constitue ainsi le 3^{ème} alphabet de la vie (Gabiuis, 2018).

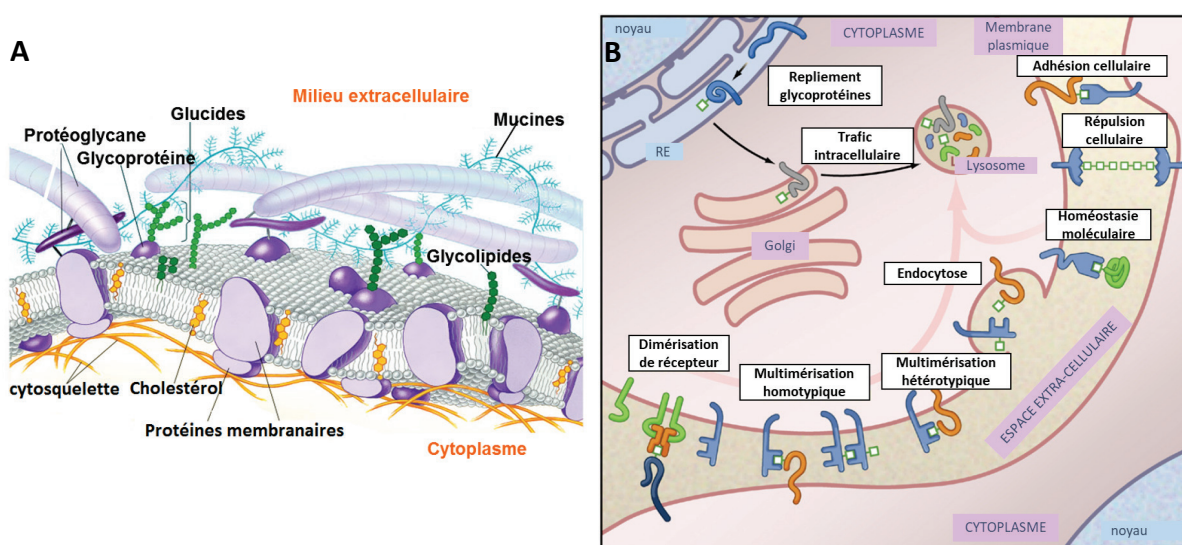


Figure I-2. Représentation de la surface cellulaire et des glycoconjugués présents (A) et description des rôles des sucres dans les cellules (B).

A. D'après Copyright© 2002 Pearson Education, Inc., publishing as Benjamin Cummings ; B. traduit de (Ohtsubo and Marth, 2006).

1.2 La glycosylation des protéines

La glycosylation est la modification post-traductionnelle la plus diverse. Contrairement aux acides nucléiques et aux protéines, la structure des glycanes ne dépend d'aucun modèle et

n'est pas prédéterminée dans le génome. La glycosylation est le résultat d'un mécanisme complexe habilement régulé par des enzymes, et dont la variabilité est influencée par des facteurs génétiques et environnementaux comme l'expression des gènes (notamment ceux qui codent pour les enzymes de la glycosylation), le métabolisme cellulaire et la structure des polypeptides. La glycosylation résulte ainsi en un panel de structures saccharidiques hétérogènes dont le nombre, la nature et les liaisons entre les monosaccharides varient (Hart and Copeland, 2010).

Les glycosylations des protéines sont classées selon la nature de l'interaction entre le sucre et la protéine. On parle ainsi de N-glycosylation, O-glycosylation et C-glycosylation, les deux premières étant les formes majoritaires qui seront détaillées ici. La glycosylation des protéines a un effet sur leur structure à travers des interactions intra et inter-moléculaires, et sur leur fonction en leur conférant de nouvelles propriétés de reconnaissance (cellule/cellule ou cellule/protéine) ou de transduction de signaux dans le cas de la fixation sur un récepteur (Opdenakker et al., 1993).

1.2.1 La N-glycosylation des eucaryotes

Chez les eucaryotes, les N-glycanes interagissent généralement avec l'azote de la chaîne latérale des résidus asparagine lorsque ceux-ci se trouvent au sein de la séquence consensus Asn-Xaa-Ser/Thr (avec Xaa désignant n'importe quel acide aminé sauf la proline) (Schmaltz et al., 2011). Environ 70 % des protéines contiennent ce séquen et 70 % des séquons sont liés à des N-glycanes (Stanley et al., 2015). Les N-glycanes sont caractérisés par un noyau pentasaccharidique commun : $\text{Man}\alpha 1-3(\text{Man}\alpha 1-6)\text{Man}\beta 1-4\text{GlcNAc}\beta 1-4\text{GlcNAc-Asn}$ (Kornfeld and Kornfeld, 1985; Kornfeld et al., 1978; Molinari, 2007). Ils sont ensuite classés en trois catégories selon la composition des monosaccharides liés à ce noyau : les N-glycanes de type oligomannose ont un noyau uniquement lié à des résidus mannose ; les N-glycanes de type complexe ont un noyau lié à plusieurs antennes commençant par un GlcNAc et composées de sucres variés tel que le galactose, le GalNAc, l'acide sialique ou le fucose (un fucose peut également être ajouté sur le premier GlcNAc du noyau) ; les N-glycanes de type hybride sont un mixte des deux précédents avec une extension de la branche $\alpha 1-6$ par des mannoses et une extension de la branche $\alpha 1-3$ par des antennes de type complexe (Stanley et al., 2015) (*Figure I-3*).

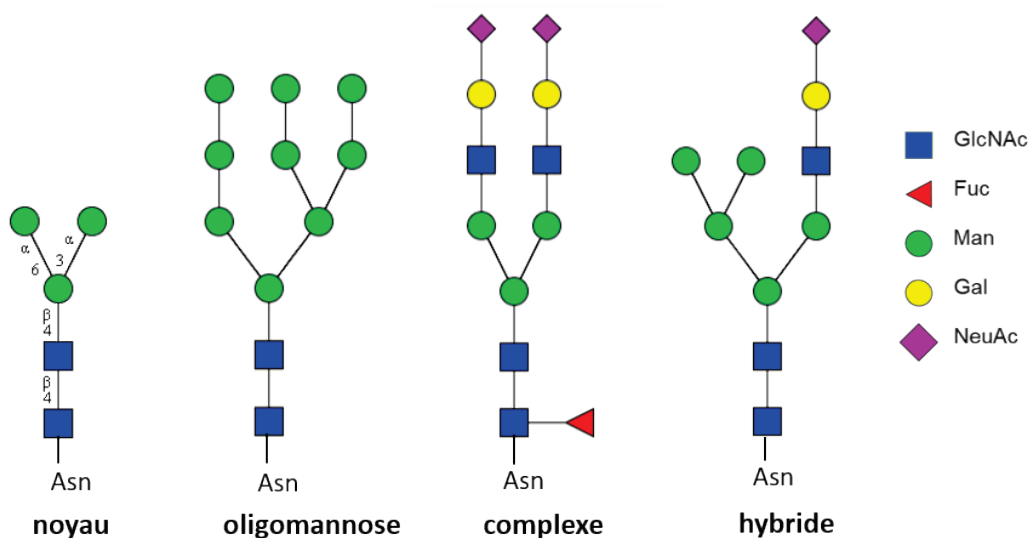


Figure I-3. Représentation schématique du noyau des N-glycanes et des trois types de N-glycanes chez l'homme.

La légende des sucres est indiquée à droite. NeuAC = acide N-acétylneuraminique (acide sialique). Les schémas ont été réalisés avec GlycoWork Bench (Ceroni et al., 2008).

Chez les Eucaryotes, les étapes de glycosylation sont similaires dans le réticulum endoplasmique (RE). Elles débutent par la formation d'un précurseur constitué d'un oligosaccharide lié à un dolichol-pyrophosphate. Cet oligosaccharide est constitué de 14 monosaccharides au moment de son transfert sur la chaîne polypeptidique en cours de synthèse, par des oligosaccharides transférases (**Figure I-4**). Suite à ce couplage, les glucoses terminaux sont enlevés par des glucosidases du RE. La protéine repliée est ensuite transférée dans l'appareil de Golgi où sont réalisées différentes étapes de glycosylation, propres à chaque espèce (Helenius and Aebi, 2001).

Chez les mammifères, on retrouve ainsi typiquement des glycanes de type complexe, dont le noyau peut être fucosylé en $\alpha 1,6$ sur le premier GlcNAc, c'est ce qu'on appelle la core-fucosylation. Chez les insectes, on retrouve des glycanes assez courts en raison de la présence de mannosidases et N-acétylglucosaminidases pendant les étapes de glycosylation. Le noyau peut à la fois être core-fucosylé en $\alpha 1,3$ et en $\alpha 1,6$. Chez les plantes, on retrouve des glycanes de type hybride et complexes dont la particularité est la présence de xylose lié en β -1,2 sur le premier mannose du noyau. Le noyau peut également être core-fucosylé en $\alpha 1,3$. Enfin, chez les champignons, les glycanes sont généralement de type oligomannose (Chung et al., 2017).

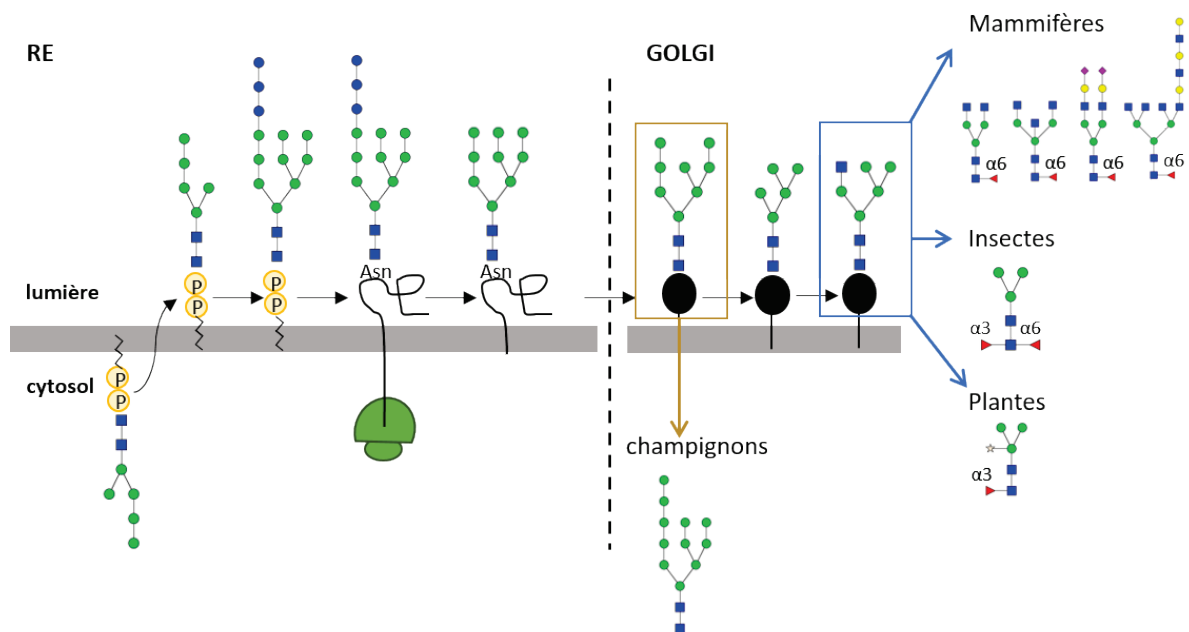


Figure I-4. Schéma simplifié des grandes étapes de la N-glycosylation selon les espèces. Les étapes dans le réticulum endoplasmique sont communes aux mammifères, insectes et plantes. Les sucres sont représentés selon la nomenclature SNFG. Schéma inspiré de (Chung et al., 2017; Helenius and Aebi, 2001).

1.2.2 La O-glycosylation

La O-glycosylation est beaucoup plus difficile à analyser que la N-glycosylation, car il y a une plus grande diversité de noyaux et elle peut être retrouvée sur n'importe quel acide aminé contenant un hydroxyle même s'il y a une nette préférence pour les sérines ou les thréonines. De plus, aucune séquence consensus spécifique n'a été identifiée pour la O-glycosylation (Schmaltz et al., 2011). Plusieurs types de O-glycosylation existent et sont classés selon la nature du premier monosaccharide ajouté sur la protéine. La plus courante est la O-glycosylation de type mucine qui consiste en l'ajout d'un GalNAc sur une sérine ou une thréonine. Les autres types de O-glycosylation consistent en l'ajout d'un GlcNAc, d'un fucose, d'un mannose ou encore d'un glucose.

Dans le cas du type O-GalNAc, la liaison du premier GalNAc (GalNAc α 1-Ser/Thr) forme l'antigène Tn. L'ajout d'un galactose lié en β 1-3 sur ce GalNAc forme ensuite le premier noyau appelé antigène T ou TF (Thomsen-Friedenreich). En tout, huit structures de noyaux différentes ont été identifiées (Steen et al., 1998). L'élongation de ces noyaux, qui a lieu dans l'appareil de Golgi, peut alors conduire à des structures très variées suite à l'ajout d'acide sialique, de fucose, de mannose ou encore de galactose. Quelques exemples de structures de O-glycanes sont représentés dans la **Figure I-5**.

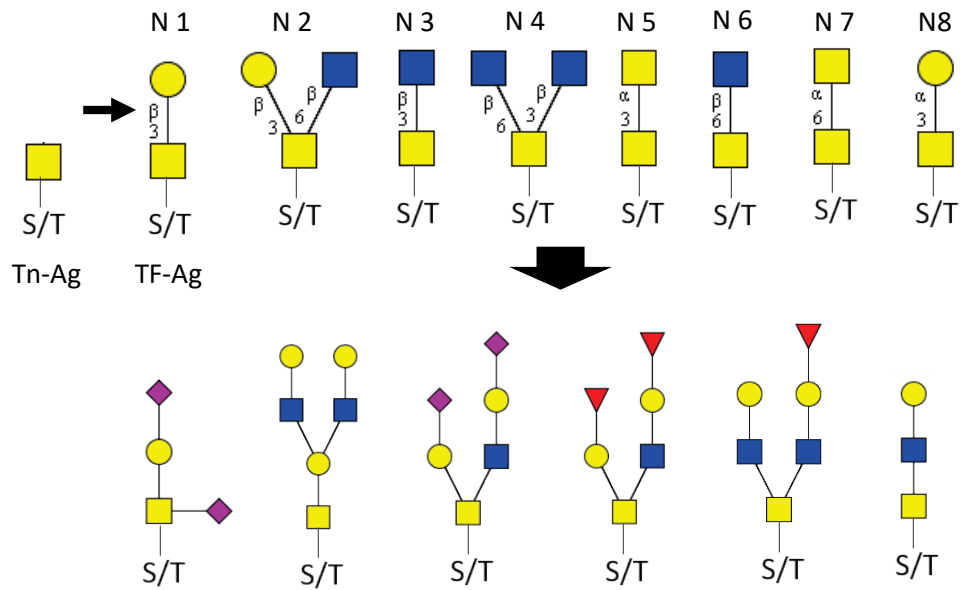


Figure I-5. Description des huit noyaux de la O-glycosylation de type mucine et de quelques exemples d'élongations possibles.

Les huit noyaux sont désignés par N1 à 8. Les antigènes Tn et TF sont nommés Tn-Ag et TF-Ag respectivement. En dessous, sont représentés les exemples possibles d'élongations.

1.2.3 Les glycosyltransférases

Les glycosyltransférases (GTs) sont des enzymes qui catalysent le transfert d'un sucre d'un donneur vers un accepteur et elles sont donc impliquées dans les différentes étapes de la glycosylation. Les sucres donneurs sont le plus souvent activés sous la forme de nucléotide-sucres (UDP Gal, GDG Man, par exemple) et les accepteurs sont généralement des sucres mais aussi parfois des lipides, des peptides ou des petites molécules. Les GTs sont nommées en fonction du sucre qu'elles transfèrent et de la nature de la liaison qu'elles réalisent sur l'accepteur (Ünligil and Rini, 2000). Elles sont répertoriées dans une base de données appelée CAZY (Carbohydrate Active Enzyme database ; Lombard et al., 2014 ; <http://www.cazy.org/>) où elles sont déclinées, à ce jour, en 107 familles en fonction de leurs similarités de séquence.

La réaction de transfert réalisée par ces enzymes peut se faire selon deux voies distinctes qui résultent soit en une inversion soit en une rétention de la configuration anomérique du sucre transféré. De fait, lorsque l'on a une réaction avec inversion, le sucre lié en α sur le donneur sera lié en β sur l'accepteur (et inversement) alors qu'il n'y a pas de changement d'anomérisation lors d'une réaction avec rétention (**Figure I-6** ; (Chang et al., 2011)).

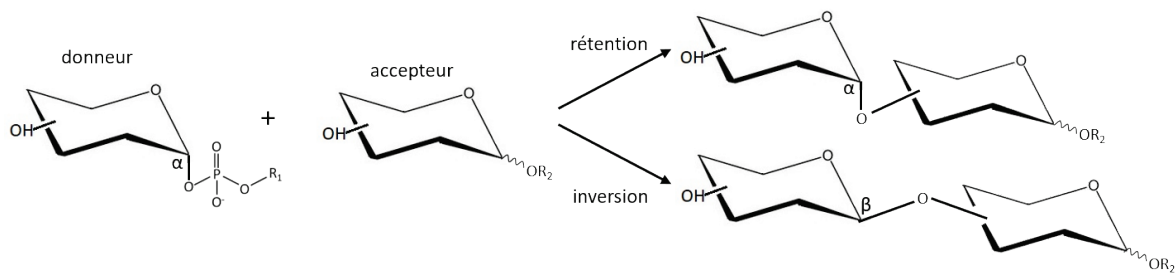


Figure I-6. Réactions des glycosyltransférases impliquant l'inversion ou la rétention du centre anomérique du sucre donneur.

Les structures des GTs se déclinent principalement en deux types de repliement appelés GT-A et GT-B (**Figure I-7**; Lairson et al., 2008) ce qui rend compte de la remarquable conservation de leur structure tridimensionnelle. Le repliement de type GT-A consiste en un domaine sandwich $\alpha/\beta/\alpha$ conduisant à la formation d'un feuillet β central continu entouré par des hélices α où seul le 6^{ème} brin β est antiparallèle. Il comprend aussi un petit feuillet β et un motif DxD permettant la fixation d'un ion divalent indispensable à l'activité de l'enzyme. Le repliement de type GT-B consiste en deux domaines Rossmann connectés par une boucle de liaison et composés d'un feuillet β de 6 à 7 brins parallèles reliés par des hélices α . Le site actif est situé dans l'espace existant entre les deux domaines (Chang et al., 2011).

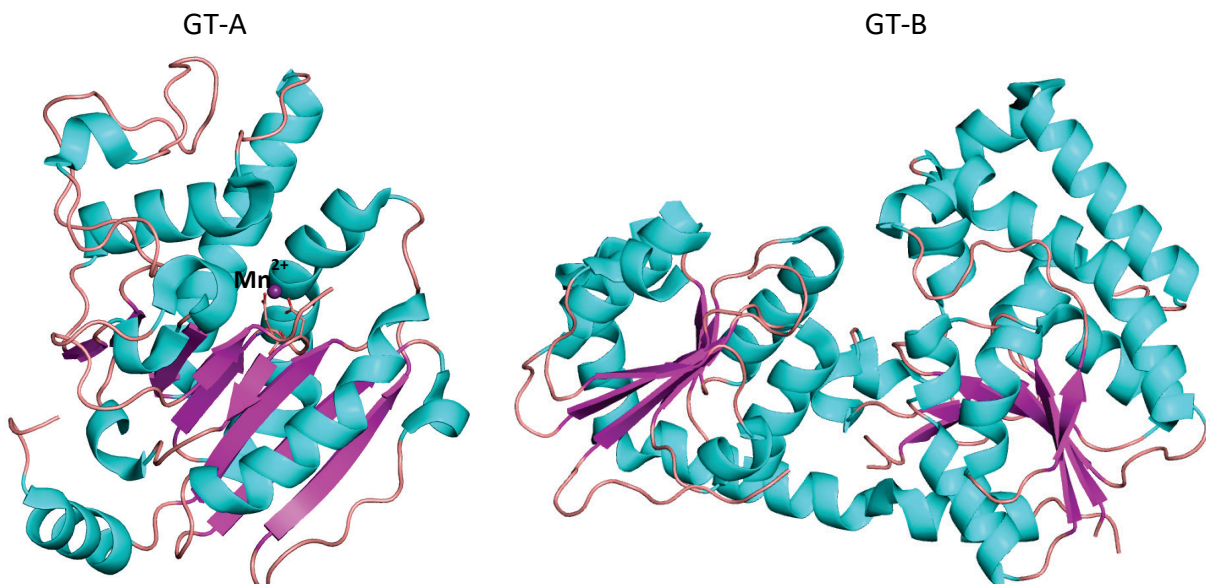


Figure I-7. Structures de glycosyltransférases de types A et B.

Le repliement de type GT-A est représenté par la glycosyltransférase A des groupes sanguins humains (pdb 3IOJ) avec la représentation de l'ion Mn^{2+} au niveau du motif DXD. Le repliement de type GT-B est représenté par l'oléandomycine glycosyltransférase OleD (pdb 2IYF).

1.3 Les anomalies de la glycosylation

De nombreuses études ont démontré l'importance du glycome dans le bon fonctionnement des organismes vivants. Grâce aux progrès technologiques qui ont permis d'améliorer la compréhension de son hétérogénéité et de sa complexité, on est désormais en mesure d'évaluer l'état de santé d'une cellule en analysant les glycanes sécrétés à sa surface. En effet, les modifications de la glycosylation peuvent entraîner des réponses inflammatoires, permettre l'entrée de virus, favoriser les métastases ou encore réguler l'apoptose (Reily et al., 2019). Les anomalies de la glycosylation peuvent résulter d'un état pathologique ou bien être héréditaires.

1.3.1 Les maladies héréditaires liées aux défauts de glycosylation

L'hypoglycosylation est retrouvée dans des maladies génétiques résultantes de défaut de glycosylation touchant aussi bien la N-glycosylation que la O-glycosylation. Parmi les maladies associées, on trouve la galactosémie, causée par une diminution de l'activité de la galactose-1-phosphate uridylyltransférase (GALT) qui permet la production d'UDP-Gal. Il s'agit d'un trouble métabolique mortel pour les nouveaux-nés qui cause une déshydratation, des diarrhées, de l'hypoglycémie et des troubles de la croissance. La galactosémie est associée à une hypoglycosylation de la N- et de la O-glycosylation des protéines, avec la prédominance de structures riches en mannoses et de glycanes tronqués ainsi qu'une forte proportion d'antigène Tn (Liu et al., 2012). Une autre maladie associée à l'hypoglycosylation est l'intolérance héréditaire au fructose (HFI : Hereditary fructosemia intolerance) causée par des dysfonctionnements de la fructose 1-6-bisphosphate aldolase B qui entraîne une accumulation de fructose-1-phosphate (F1P). Cela est associé à une forte prédominance d'antigènes TF (Michelakakis et al., 2009).

Parmi les troubles héréditaires de la glycosylation, on retrouve l'ensemble des anomalies congénitales de la glycosylation (ou CDG pour « congenital disorders of glycosylation »). La majorité des CDG sont des maladies autosomiques récessives (Jaeken and Péanne, 2017). Ces maladies peuvent entraîner de graves complications neurologiques, hématologiques, gastro-intestinales, rénales, cardiovasculaires, ou encore ophtalmologiques, et sont à l'origine d'importants retards de développements intra-utérins (Al Teneiji et al., 2017) mais peuvent aussi avoir des formes moins sévères avec des phénotypes adultes presque normaux (Jaeken and Péanne, 2017). Elles affectent les processus de glycosylation à de nombreux niveaux et sont très variées puisque plus de 100 types de CDG différents ont été décrits (Hennet and

Cabalzar, 2015). La plupart des patients diagnostiqués avec une CDG présentent une carence en phosphomannomutase (PMM) (Van Schaftingen and Jaeken, 1995). Cette enzyme est responsable de la conversion du Man-6-phosphate en Man-1-phosphate et cette carence entraîne donc une faible production de GDP-Man et de Dol-P-Man, deux précurseurs de la glycosylation.

1.3.2 Altérations de la glycosylation dans les cancers

1.3.2.1 Généralités

Les études ont montré que les aberrations de la glycosylation sont retrouvées dans tous types de cancers. Elles ne sont généralement pas directement à l'origine du processus de cancérisation des cellules mais sont le résultat d'une initiation oncogénique. Par contre, les altérations de la glycosylation sont des événements clés qui peuvent favoriser le grossissement des tumeurs, les processus d'invasion et les métastases (Hakomori, 2002 (*Figure I-8*)). En effet, le grossissement des tumeurs dépend de la capacité des cellules cancéreuses à échapper aux contrôles de la division cellulaire, aux signaux de mort cellulaire et à la surveillance immunitaire et à migrer vers des sites métastatiques.

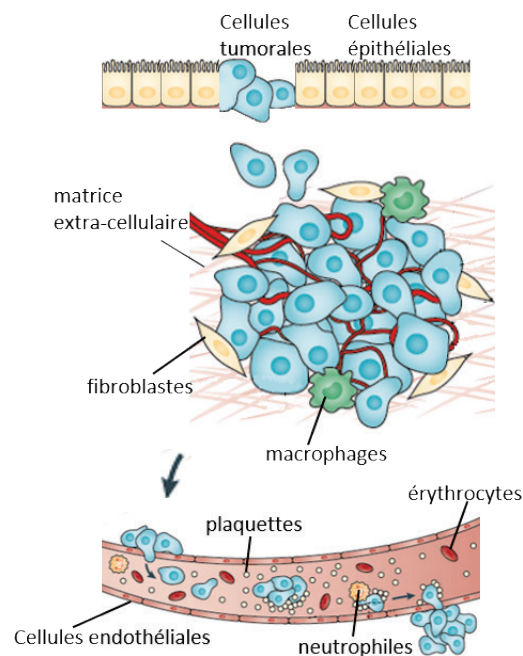


Figure I-8. Dissociation des cellules tumorales et invasion conduisant à la formation de tumeurs métastatiques d'après (Pinho and Reis, 2015).

Les glycanes présents en surface des cellules tumorales favorisent l'invasion des tumeurs en modifiant les processus de signalisation et les interactions cellule-cellule ou cellule-matrice extra-cellulaire.

La glycosylation jouant un rôle important dans ces processus, ses altérations peuvent favoriser les processus de cancérisation (Reily et al., 2019). Par exemple, les motifs de glycosylation présents sur les ligands et récepteurs de la signalisation des facteurs de croissance peuvent avoir un rôle important dans le développement de cancers (Gomes Ferreira et al., 2018). Les aberrations de la glycosylation des glycoprotéines peuvent être dues à plusieurs facteurs tels que la sous- ou sur-expression de glycosyltransférases, une altération de l'activité des glycosidases, une mauvaise conformation tertiaire du peptide, ou encore à des carences en substrats ou à une mauvaise localisation des glycosyltransférases dans l'appareil de Golgi (Pinho and Reis, 2015).

Les motifs de glycosylation ont été les premiers biomarqueurs de cancer et sont toujours utilisés pour identifier les cellules cancéreuses (Schultz et al., 2013). Les altérations de glycosylation généralement observées dans les cellules cancéreuses sont l'augmentation des structures Lewis sialylées et de glycanes polysialylés, des taux de core-fucosylation anormaux, des altérations et des troncations des branchements des N-glycanes ou des O-glycanes de type mucine avec l'exposition des antigènes T et TF (Reily et al., 2019; Satomaa et al., 2009 (**Figure I-9**)). Un grand nombre de ces motifs de glycosylation sont appelés « oncofœtaux » car ils sont habituellement retrouvés dans les premières phases du développement (Pinho and Reis, 2015).

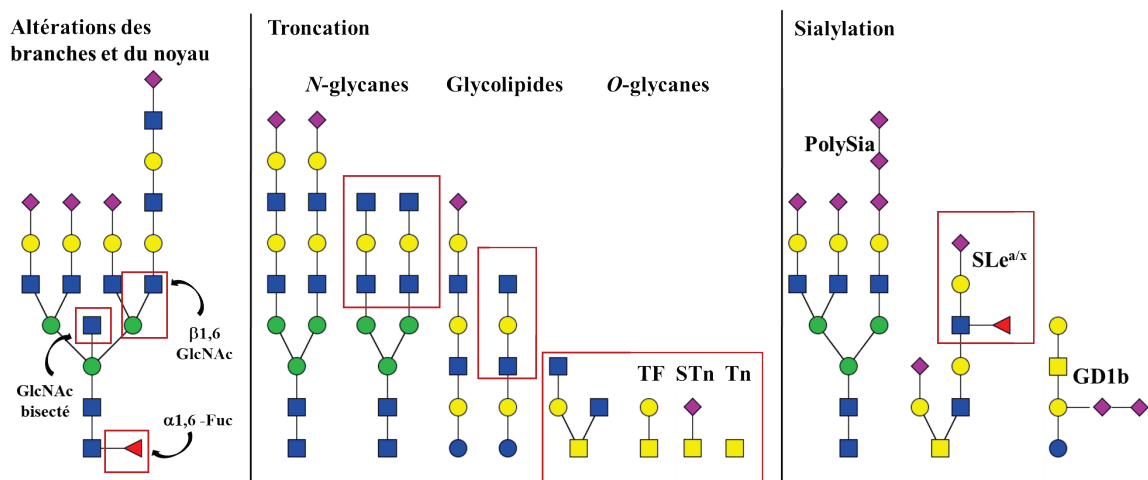


Figure I-9. Représentation de motifs de N- et O- glycosylation altérés retrouvés dans les cancers.

Les motifs de glycosylation associés aux processus de cancérisation sont appelés TACA (Tumor Associated Carbohydrate Antigen). Au fil des années, les études sur le sujet ont permis de mieux comprendre l'implication de ces structures dans le développement des

cancers. La sialylation des glycanes joue un rôle important dans la reconnaissance, l'adhésion et la signalisation cellulaire et l'augmentation de ces taux est souvent associée aux cancers. Des structures polysialylées sont observées dans des formes agressives de cancers du poumon, de neuroblastomes et de gliomes (Falconer et al., 2012). De même, les structures Lewis^a et Lewis^x sialylées (SLe^x et SLe^a) sont retrouvées à des taux élevés chez les patients cancéreux ayant de faibles chances de survie. Les structures SLe^x sont reconnues par des sélectines, des molécules d'adhésion des cellules vasculaires. Dans les cancers, l'interaction des sélectines avec les structures SLe^x forme des embolies de cellules cancéreuses et de plaquettes qui favorisent leur stabilisation sur les cellules endothéliales et donc la formation de métastases (Pinho and Reis, 2015). Les structures SLe^a sont utilisées en pratique clinique comme marqueur pour le diagnostic de cancers pancréatiques, colorectaux, gastriques ou biliaires (Locker et al., 2006; Reis et al., 2010). La transformation maligne est également associée à une augmentation de l'expression des N-glycanes complexes possédant un branchement β 1,6 généralement lié à une structure polylactosamine reconnue par les galectines. Celles-ci modulent divers processus cellulaires comme en particulier l'adhérence intercellulaire et jouent donc un rôle important dans les cancers (Crocì et al., 2014). Au contraire, l'augmentation du taux de GlcNAc bisecté est favorable à la suppression des tumeurs et des métastases. Enfin, la troncation des O-glycanes, conduisant à l'exposition des antigènes Tn, SialylTn et TF est un phénomène commun dans de nombreuses tumeurs. Elle résulte de l'altération de l'expression des GalNAc transférases et conduit à une diminution de l'adhésion cellulaire, une augmentation de la croissance des tumeurs, et à leur invasion (Pinho and Reis, 2015).

Les connaissances sur les antigènes carbohydrates associés aux tumeurs et l'étude de leur présence en surface des cellules cancéreuses et sur les glycoprotéines du sérum de patients, permettent de les utiliser comme biomarqueurs diagnostiques, d'évaluer le pronostic des patients et de suivre les réponses aux traitements. Les études montrent cependant que l'utilisation d'un seul biomarqueur est souvent peu fiable et que la détermination du profil complet du glycome serait beaucoup plus efficace pour évaluer la progression des maladies (Reily et al., 2019).

1.3.2.2 La core fucosylation

La synthèse de GDP-fucose peut se faire selon deux voies différentes, appelées voie *de novo* et voie de récupération (Becker and Lowe, 2003 ; **Figure I-10**). La voie *de novo*

transforme le GDP-mannose en GDP-fucose grâce aux réactions de catalyse réalisées par la GDP-mannose-4,6-déhydratase (GDM) et la GDP-fucose synthase (FX). La réaction de récupération synthétise le GDP-fucose à partir de fucose libre. Les fucosyltransférases utilisent ce GDP-fucose comme substrat pour fucosyler des glycoprotéines par exemple. C'est le cas de la fucosyltransférase 8 (Fut8) qui catalyse la liaison d'un fucose en $\alpha 1,6$ sur le premier résidu GlcNAc des N-glycanes. La Fut8 est la seule enzyme capable de réaliser ce type de liaison (Miyoshi et al., 2008). La core-fucosylation est le type de fucosylation retrouvé le plus communément dans les glycoprotéines du sérum humain (Lin et al., 2011).

La core-fucosylation joue des rôles importants dans les cellules et elle est notamment impliquée dans la croissance et la différenciation. Par exemple, il a été montré que la core-fucosylation des N-glycanes est requise pour la fixation du facteur de croissance épidermique, EGF, à son récepteur (EGFR) (Wang et al., 2006). De même, les diminutions de la core-fucosylation du récepteur TGF- $\beta 1$ (Wang et al., 2005) ou de l'intégrine $\alpha 3\beta 1$ (Zhao et al., 2006) résultent en une diminution de leur affinité pour leurs ligands et donc en une inhibition de voies de signalisation correspondantes. L'extinction du gène de la Fut8 conduit ainsi à des retards de croissance, à des décès prématurés au cours du développement postnatal, et à des emphysèmes pulmonaux (destruction des parois des alvéoles) (Wang et al., 2005).

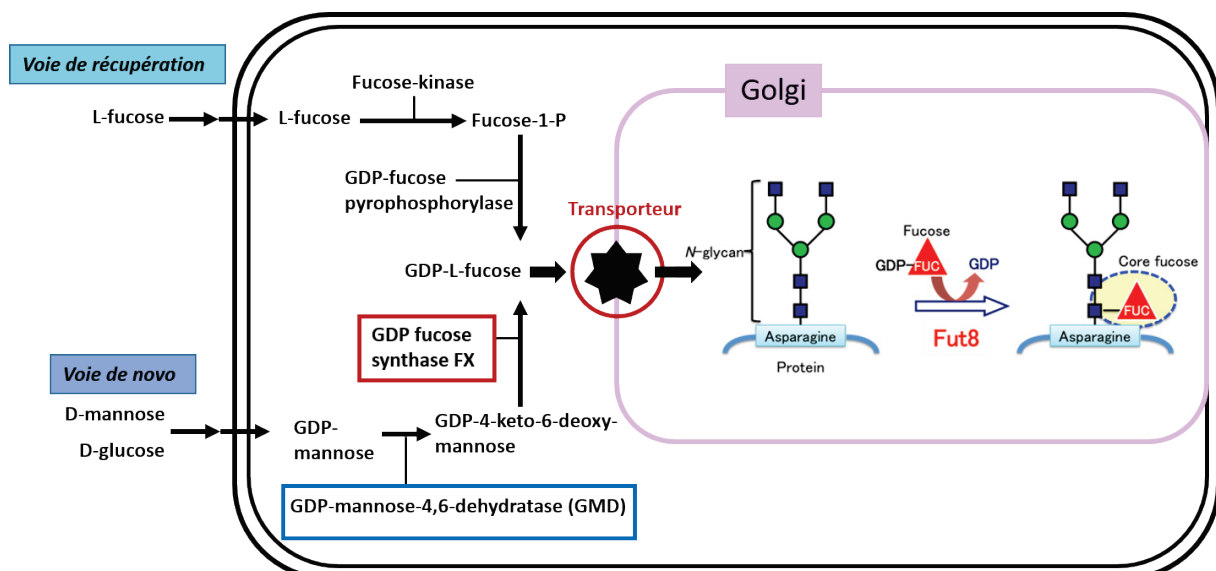


Figure I-10. Voies de synthèse du GDP fucose et réaction de core-fucosylation en $\alpha 1,6$ des N-glycanes par la fucosyltransférase 8 (Fut8) dans l'appareil de Golgi.

Figure inspirée de (Miyoshi et al., 2008). L'utilisation du GDP-fucose est présentée ici uniquement pour la fucosyltransférase 8 qui catalyse la liaison en $\alpha 1,6$ sur les N-glycanes mais d'autres fucosyltransférases utilisent le GDP-fucose dans les cellules.

Les taux anormaux de la core-fucosylation sont associés à de nombreuses maladies telles que les cancers où elle est souvent augmentée comme pour la molécule d'adhésion L1CAM dans les mélanomes métastatiques (Agrawal et al., 2017). En effet, L1CAM est un médiateur de l'expansion et de la croissance métastatique et la core-fucosylation de cette protéine limite son clivage protéolytique par la plasmine, ce qui favorise l'invasion cellulaire. L'augmentation de la core-fucosylation du transporteur de cuivre CTR1 contribue à la progression des cancers épithéliaux ovariens (Lv et al., 2019). La core-fucosylation de la protéine PD-1 régule, elle, l'activation des cellules T qui font partie de la réponse immunitaire anti-tumorales (Okada et al., 2017). Plusieurs glycoprotéines du sérum voient leur taux de core-fucosylation augmenter au cours du processus cancéreux comme l' α -fétoprotéine (AFP-L3 ; Nakagawa et al., 2008), l'antigène spécifique de la prostate (PSA ; Saldova et al., 2011) et l'haptoglobine (Hpt ; Lin et al., 2011; Okuyama et al., 2006) qui sont utilisées comme biomarqueurs pour la détection de l'hépatocarcinome, et des cancers de la prostate et du pancréas respectivement. La glycoforme 3 de l'AFP, ou AFP-L3, est d'ailleurs un biomarqueur approuvé pour la détection précoce de carcinomes hépatocellulaires (HCC) (Sato et al., 1993). Récemment, une étude de la glycosylation de tissus du carcinome épidermoïde du poumon (LUSC) a aussi montré une augmentation de la core fucosylation comparée aux tissus normaux (Ma et al., 2019). Il existe aussi des exemples de cancer associés à une diminution du taux de core-fucosylation global dans le sérum et les tissus atteints comme dans le cancer de l'estomac (Zhao et al., 2014).

1.3.2.3 L'antigène TF (Thomsen–Friedenreich)

L'antigène TF, aussi appelé antigène T, est une structure intermédiaire de la biosynthèse des O-glycanes de type mucine. Il correspond au noyau de type 1 de la O-glycosylation (Gal β 1-3GalNAc α 1-Ser/Thr). La O-glycosylation débute par l'ajout d'un GalNAc sur une sérine ou une thréonine par des N-acétylgalactosaminyltransférases (GALNTs). Ce GalNAc est ensuite galactosylé par la GalNAc3 β -galactosyltransférase 1 (C1GalT1) avec l'aide de sa chaperone Cosmc (C1GALT1C1) localisée dans le réticulum endoplasmique (Hung et al., 2014). Dans les tissus normaux, l'antigène TF est ensuite modifié par d'autres glycosyltransférases pour générer des O-glycanes plus complexes (Tian and Ten Hagen, 2009).

Dans la plupart des tissus sains, l'antigène TF en surface des cellules est masqué et n'est donc pas immunoréactif (Springer, 1984). Il est cependant fortement exprimé, exposé et donc immunoréactif dans de nombreux types de tumeurs et en particulier dans les tumeurs

épithéliales. On le retrouve dans environ 100 % des cancers ovariens, 85 % des cancers du sein, 60 % des cancers du côlon, 50% des leucémies et 15% des cancers rénaux (Karsten and Goletz, 2015). L'antigène TF est impliqué dans la réponse immunitaire et est notamment retrouvé sur les glycoprotéines CD4 et CD8 des cellules T (Butcher et al., 2014). Dans le cancer du côlon, les mucines et notamment MUC-1 présentent généralement l'antigène TF qui sert de ligand pour la lectine Galectine-3 (Gal-3) qui est, en outre, impliquée dans les processus métastatiques (Bresalier et al., 1996). L'interaction de Gal-3 avec MUC-1 via l'antigène TF est un principe central de l'adhésion métastatique des cellules (Yu et al., 2007). Dans les cancers du sein métastatiques, des niveaux élevés d'antigène TF et de Gal-3 sont aussi mis en évidence (Glinsky et al., 2000). L'antigène TF représente donc un bon marqueur pour la détection de cancers et le suivi de leur progression afin d'évaluer particulièrement le risque de formation de métastases.

Les mécanismes qui conduisent à la surexpression de l'antigène TF sont difficiles à déterminer. Les dernières études ont suggéré que l'expression de l'antigène TF pouvait être liée à la dissipation du gradient de pH dans l'appareil de Golgi (Karsten and Goletz, 2015). L'alcalinisation de la lumière de l'appareil de Golgi est en effet fréquemment observée dans les cellules tumorales, ce qui entraîne des modifications de localisation de certaines glycosyltransférases. Cela peut donc conduire à des O-glycosylations incomplètes exposant l'antigène TF (Rivinoja et al., 2006).

1.3.3 Glycosylation et efficacité de certains anticorps thérapeutiques

Les progrès rapides dans l'ingénierie des anticorps ont permis de les utiliser dans de nombreuses applications biopharmaceutiques pour le traitement des maladies autoimmunes, des cancers et de la septicémie. Plus de 50 anticorps monoclonaux (mAbs) ont été approuvés aux États-Unis et en Europe en tant que médicament pour le traitement de cancers et de maladies chroniques (Ecker et al., 2014).

Les anticorps monoclonaux (mAbs) peuvent cibler les tumeurs grâce à la reconnaissance spécifique d'antigènes et en recrutant des effecteurs tels que les macrophages, les cellules dendritiques, les cellules natural killer (NK), les cellules T et les composants de la voie du complément (Jefferis et al., 1998). Ces recrutements se font par l'interaction du fragment Fc des immunoglobulines gamma (IgG) avec les récepteurs de cellules immunitaires (FcγRs) ou la protéine du complément C1q. Il existe également les récepteurs néonataux (FcRn) qui prolongent la demi-vie des mAbs (Liu, 2015). Les fonctions effectrices des anticorps sont par

exemple la cytotoxicité à médiation cellulaire dépendante des anticorps (ADCC pour Antibody-dependent cell-mediated cytotoxicity), la phagocytose dépendant des anticorps (ADCP pour antibody-dependant cell-mediated phagocytosis) et la lyse cellulaire complément-dépendante (CDC pour complement-dependant-cytotoxicity) (Liu, 2015 ; **Figure I-11**).

Tout comme les IgG naturelles, les anticorps monoclonaux (mAbs) recombinants thérapeutiques sont glycosylés. L'altération de leur composition en glycanes peut entraîner des changements conformationnels du domaine Fc, ce qui peut modifier leur affinité pour les récepteurs Fcγ et donc modifier leurs fonctions effectrices (Liu, 2015). Les N-glycanes des IgG humaines ont généralement des structures complexes, constituées du noyau des N-glycanes sur lequel peuvent se greffer un core-fucose, des galactoses ou des acides sialiques.

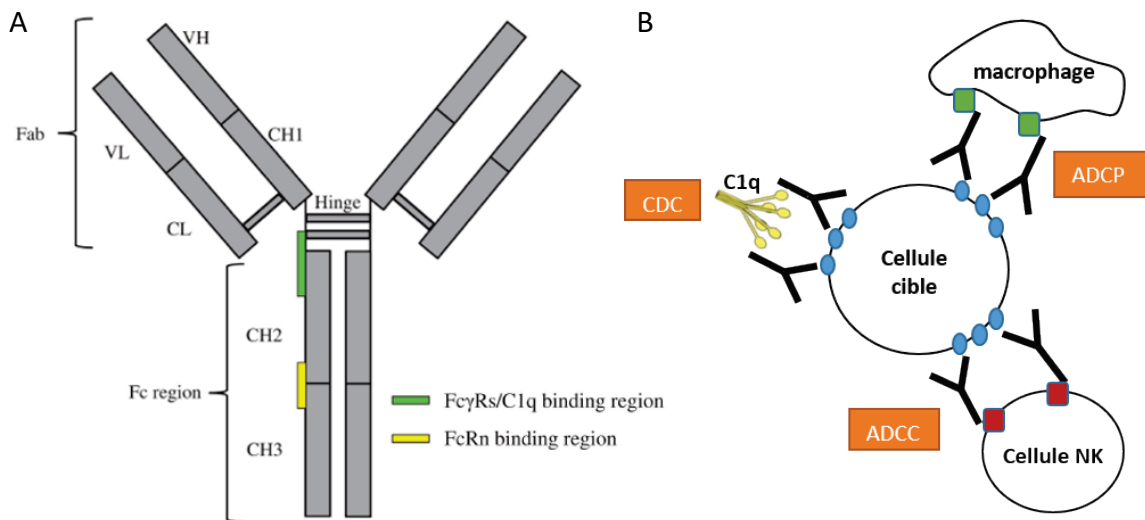


Figure I-11. A. Représentation schématique d'une immunoglobuline (Saxena and Wu, 2016) B. Schéma des principaux mécanismes d'action des anticorps.

A. La chaîne lourde est constituée des parties VH, CH1, Hinge, CH2 et CH3. Les chaînes légères sont constituées des parties VL et CL. Des ponts disulfures inter-chaînes permettent de stabiliser la structure. En vert, les sites d'interaction avec les FcγRs et C1q, en jaune le site d'interaction avec le FcRn. B. Les anticorps reconnaissent leurs antigènes via leurs fragments Fabs et se fixent à leurs récepteurs via leur fragment Fc (récepteurs FcγR des cellules natural killer (NK) et macrophages ou protéines C1q dans l'exemple).

Parmi les motifs de glycosylation du Fc impactant les fonctions effectrices des anticorps, la core-fucosylation joue un rôle important. La plupart des anticorps développés en tant que médicament sont des IgG1 humaines. Ces immunoglobulines sont caractérisées par une grande richesse en N-glycanes core-fucosylés et leur mécanisme d'action est principalement basé sur l'ADCC (Mori et al., 2007). Il a été montré que la core-fucosylation des N-glycanes

du Fc de certaines IgG1 diminuait l'affinité de l'interaction entre le Fc et le récepteur FcγRIIIa (Yamane-Ohnuki and Satoh, 2009). Or, le récepteur FcγRIIIa est généralement occupé par les IgG du plasma, donc si l'affinité des anticorps thérapeutiques est trop faible, il va y avoir compétition entre les IgG du plasma et les anticorps thérapeutiques pour le FcγRIIIa, ce qui va diminuer l'ADCC induite par les anticorps thérapeutiques (Preithner et al., 2006). À l'inverse, les anticorps thérapeutiques non fucosylés ont une meilleure affinité pour le récepteur FcγRIIIa que les IgG du sérum, ce qui permet d'éviter l'effet inhibiteur de ces IgG sur l'ADCC (Satoh et al., 2006).

L'objectif est donc désormais d'être capable de produire des anticorps thérapeutiques non-fucosylés. Cela est maintenant possible notamment en utilisant des cellules knockout pour le gène FUT8 responsable de la core-fucosylation (Satoh et al., 2006). Néanmoins, il est important d'avoir des outils pour le contrôle-qualité des anticorps produits afin de pouvoir vérifier facilement et rapidement les taux de core-fucosylation lors du processus de production et pour la sélection des clones adéquats.

2. Les lectines : outils performants d'analyse de la glycosylation

2.1 Généralités sur les lectines

Les lectines sont des protéines ubiquitaires d'origine non immunitaire qui se lient de manière spécifique et réversible aux carbohydrates sans les modifier. Elles n'ont pas d'activité catalytique et beaucoup présentent des propriétés d'agglutination d'où leur dénomination fréquente comme agglutinine ou hémagglutinine (Lis and Sharon, 1998).

Les lectines ont été initialement découvertes chez les plantes avec la première description pour la ricine datant de 1888 dans la thèse de Peter Hermann Stillmark. Cette lectine extraite à partir de ricin et extrêmement toxique, a permis notamment de démontrer la spécificité des anticorps vis-à-vis d'un antigène donné, le phénomène de mémoire immunitaire et le transfert de l'immunité de la mère à l'enfant (Sharon and Lis, 2004). Dans les années 1940, William C. Boyd et Karl O. Renkonen ont découvert la spécificité de certaines hémagglutinines pour les groupes sanguins. En 1954, la capacité des agglutinines de plantes à faire la distinction entre les érythrocytes de différents groupes sanguins a mené Boyd et Shapleigh à proposer le terme « lectine » pour ces protéines. Ce terme est issu du latin « legere » signifiant « lire ». Il a été par la suite généralisé pour inclure toutes les agglutinines spécifiques de sucres n'ayant pas d'origine immune (Sharon and Lis, 1972).

Les premières lectines ayant été découvertes chez les plantes, c'est dans ces organismes que leurs rôles physiologiques ont préalablement été étudiés. La première publication à ce sujet date de 1976 et a permis de mettre en avant le rôle insecticide des lectines, notamment pour la protection des légumes contre des insectes prédateurs de graines (Janzen et al., 1976). Il a également été observé que les lectines WGA (Wheat Germ Agglutinin), PNA (Peanut Agglutinin) et SBA (Soybean Agglutinin) inhibaient la sporulation et la croissance de champignons tel que *Trichoderma viride*, *Penicilium notatum* et *Aspergillus niger* (Barkai-Golan et al., 1978), ce qui a confirmé le rôle protecteur des lectines chez les plantes. Elles sont également capables de réaliser des symbioses avec des bactéries, par exemple dans le but de fixer l'azote atmosphérique. Les lectines ont des rôles variés au sein des plantes via leur interactions avec des protéines de stockage et des enzymes (Rüdiger and Gabius, 2001). Chez les virus et les bactéries, les lectines favorisent les processus d'infection de l'hôte (Taylor et al., 2015). Chez les animaux, elles ont des fonctions très variées tel que l'acheminement des glycoconjugués dans les cellules, des fonctions de chaperonne, la médiation de l'endocytose, la croissance cellulaire, les interactions cellulaires et elle jouent un rôle majeur dans le fonctionnement du système immunitaire notamment dans la reconnaissance du non-soi et dans les mécanismes de défense (Kilpatrick, 2002). Les rôles des lectines les plus communs sont présentés dans la **Figure I-12**.

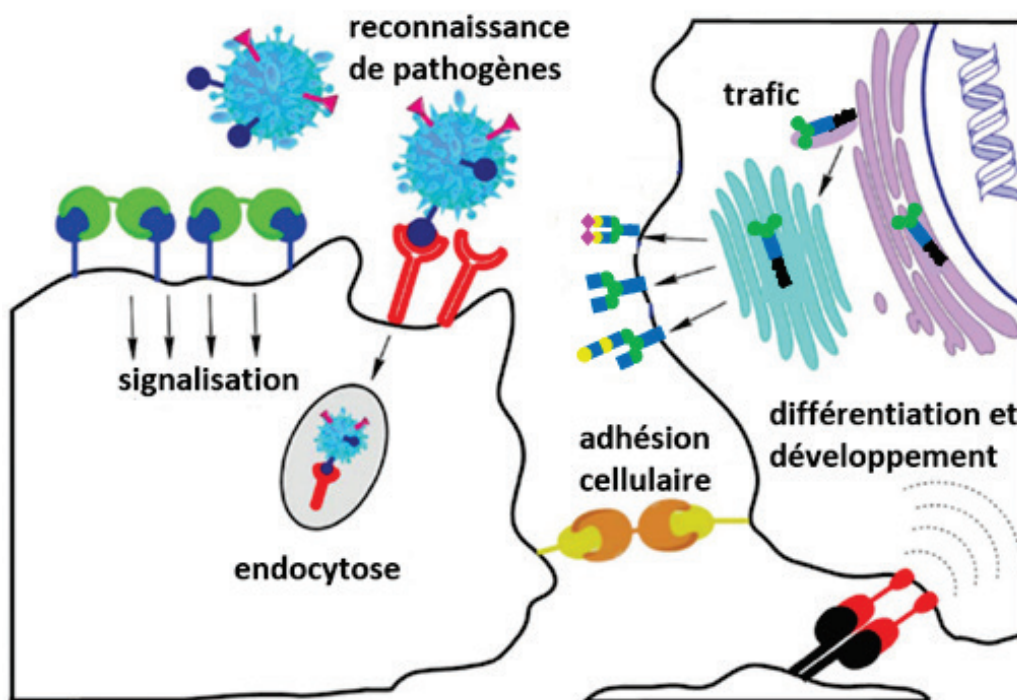


Figure I-12. Principaux rôles des lectines *in vivo*.
D'après (Bojarová and Křen, 2016).

Les lectines sont ainsi capables de déchiffrer l'information structurale présente au sein des glycoconjugués, que l'on appelle le glycode. Les lectines reconnaissent leurs ligands via des interactions faibles mais ceci est généralement compensé grâce au phénomène de multivalence, c'est-à-dire à la fixation de plusieurs ligands. La multivalence des lectines provient soit de la répétition du motif de reconnaissance du sucre, soit d'oligomérisation ou association de plusieurs protomères entre eux ou soit d'une combinaison de ces deux dernières. Cette multivalence est à l'origine des propriétés agglutinantes des lectines et permet d'assurer une haute spécificité et d'augmenter l'affinité (Lee and Lee, 1995). Les lectines présentent une grande diversité de repliement et de structure quaternaire. L'ensemble de leur structures 3D sous forme libre ou complexée avec des ligands a été regroupé récemment dans la nouvelle base de donnée Unilectin (Bonnardel et al., 2019 ; <https://www.unilectin.eu/unilectin3D>). Elle présente une classification de l'ensemble des lectines connues en familles en s'appuyant sur leur origine, leur structure tridimensionnelle, et leur spécificité.

2.2 Les lectines fongiques

Ces dernières années, les lectines fongiques, c'est-à-dire les lectines issues de champignons, ont fortement attiré l'attention des glycobiologistes. Environ 105 lectines ont été identifiées dans des champignons (Hassan et al., 2015). Parmi celles qui sont connues, 82 % sont issues de champignons supérieurs (généralement des basidiomycètes et quelques ascomycètes), 15 % sont issues de moisissures, et 3 % sont issues de levures (organismes unicellulaires). Les premières lectines fongiques ont été découvertes au cours de recherches sur la toxicologie des champignons supérieurs. Ainsi, la première lectine fongique à avoir été décrite par Ford en 1910, a été associée à la toxicité du champignon Amanite tue-mouches (*Amanita muscaria*) duquel elle est issue (Singh et al., 2010).

Comme chez les autres organismes, les lectines fongiques ont des rôles très variés au sein des champignons. Elles servent de protéines de stockage, interviennent dans les processus de croissance et sont impliquées dans la reconnaissance de glycoconjugués en surface des cellules de l'hôte lors des processus de symbiose et d'infection. Elles jouent aussi un rôle essentiel dans la reconnaissance du non-soi donc dans l'immunité innée, seule ligne de défense des champignons. Certaines ont un rôle de défense notamment du fait de leur toxicité qui leur confère des propriétés insecticides, vermicides ou antivirales (Varrot et al., 2013). Ces propriétés permettent de les utiliser dans diverses applications, comme en glycobiologie, dans le but d'isoler des glycoconjugués spécifiques ou de suivre les modifications de glycanes

présents en surface des membranes cellulaires, dans des cas physiologiques comme pathologiques. Certaines d'entre elles ont des activités antitumorales, antiprolifératives ou mitogéniques, et ce sont aussi des outils intéressants pour l'étude de l'implication des sucres dans les maladies telles que les cancers. (Varrot et al., 2013 ; **Figure I-13**).


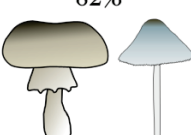
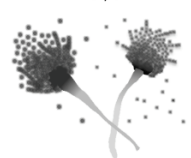

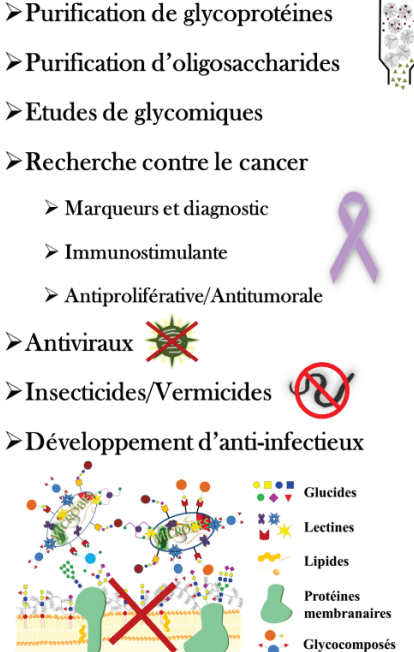
ROLES	ORIGINE	APPLICATIONS
<ul style="list-style-type: none"> ➤ Cohésion des tissus ➤ Développement et morphogénèse ➤ Reconnaissance moléculaire/interactions organismiques <ul style="list-style-type: none"> ➤ Mycorhize ➤ Défense ➤ Parasitisme ➤ Infections  <ul style="list-style-type: none"> ➤ Flocculation/Reproduction 	<p>Champignons 82%</p>  <p>Moisissures 15%</p>  <p>Levures 3%</p> 	<ul style="list-style-type: none"> ➤ Purification de glycoprotéines ➤ Purification d'oligosaccharides ➤ Etudes de glycomiques ➤ Recherche contre le cancer <ul style="list-style-type: none"> ➤ Marqueurs et diagnostic ➤ Immunostimulante ➤ Antiproliférative/Antitumorale ➤ Antiviraux ➤ Insecticides/Vermicides ➤ Développement d'anti-infectieux  <ul style="list-style-type: none"> Glucides Lectines Lipides Protéines membranaires Glycocomposés

Figure I-13. Récapitulatif des divers rôles biologiques des lectines fongiques et de leurs applications en biologie et en recherche biomédicale.
Traduit de (Varrot et al., 2013).

Récemment, un système de classification des lectines fongiques a été proposé en se basant sur les similarités structurales et les similarités de séquence. Mais à cause du manque d'informations sur ce type de lectines, leur classification est encore difficile. De nombreuses lectines fongiques sont classées dans des familles communes aux lectines de plantes ou d'animaux, tandis que d'autres sont classées dans de nouvelles familles telles que les familles AAL-like ou PVL-like par exemple (**Tableau I-1**).

Peu de structures tridimensionnelles de lectines fongiques ont été pour le moment résolues, notamment parce que les séquences primaires ne sont pas toujours connues. En 2015, les structures cristallographiques de seulement 17 lectines fongiques étaient connues. Parmi elles, différents types de repliement différents ont été identifiés dont les repliements de type β -propeller, galectin, β -trefoil et actinoporin (Hassan et al., 2015).

Familles de lectines	Spécificités
Galectin	Galactose
Ricin B-chain like (R-type)	Galactose
Jacalin-related	Mannose
Monocot (GNA-related)	Mannose
ABL-like (α/β -sandwich Actonoporin-like)	Galactose, GlcNAc, GalNAc
PVL-like (7-bladed β -propeller)	GlcNAc
AAL-like (6-bladed β -propeller)	Fucose
PhoSL-like	Fucose
LSL	LacNAc

Tableau I-1. Liste des familles dans lesquelles sont classées les lectines fongiques et des spécificités des lectines retrouvées dans chaque famille (Kobayashi and Kawagishi, 2014).

Le monde fongique est donc une source encore peu explorée et peu exploitée pour de nouvelles lectines avec un fort potentiel en tant qu'outils d'étude en glycobiologie ou pour des applications biotechnologiques ou biomédicales. Les avancées en séquençage de génome et en production de lectines sous forme recombinante devraient permettre d'apporter des avancées rapides dans ce domaine.

2.3 Utilisation des lectines comme outil de reconnaissance des sucres

Les propriétés des lectines à reconnaître des sucres de façon spécifique et de pouvoir agir comme des toxines pour certaines d'entre elles, permet de les utiliser dans de nombreuses applications, que ce soit en tant qu'outil, en tant qu'agent insecticide (Kaur et al., 2009), anti-fongique (Yan et al., 2005), antiviral (Swanson et al., 2010) ou anticancéreux (Choi et al., 2004).

2.3.1 Analyse de la glycosylation

Une des applications possibles des lectines en analyse de la glycosylation est l'utilisation de colonnes à lectines dans le but de purifier spécifiquement un type de sucre ou glycoconjugué. Le principe est simplement de fixer les lectines sur une résine, et ensuite d'y

ajouter les glycoconjugués d'intérêt, qui se fixeront, ou non, à la colonne selon la spécificité de la lectine fixée et les sucres présents sur les glycoconjugués. Une des premières colonnes à lectine a été élaborée à partir de la lectine Con A (concavaline A) en 1970 (Donnelly and Goldstein, 1970). Cela a permis de démontrer que cette lectine reconnaissait les mêmes ligands qu'en solution, ce qui permettait de séparer des polysaccharides et glycoprotéines d'un mélange complexe. De plus, les lectines interagissent avec les sucres via des interactions non-covalentes tel que les interactions hydrogènes, hydrophobes, van der Waals ou électrostatiques, et ont des affinités généralement assez faibles pour leur ligands (entre 10^{-3} et 10^{-6} M) (Dan et al., 2016). Cela permet de facilement détacher les sucres fixés, par compétition avec d'autres substances reconnues par les lectines. Actuellement plusieurs types de colonnes d'affinité à lectines sont commercialisées, c'est le cas de la colonne Con A-Sepharose et Lentil Lectin Sepharose 4B fournies par GE Healthcare Life Sciences.

Les lectines peuvent également être utilisées en histochimie et cytochimie. En effet, elles peuvent facilement être marquées avec des sondes fluorescentes, ce qui permet de repérer des sucres spécifiques sur des cellules et des tissus en utilisant des microscopes à fluorescence ou la cytométrie en flux. Elles peuvent également être marquées avec de la biotine puis conjuguées à des enzymes (tel que la peroxydase de raifort) qui permettent de transformer un substrat donné en un produit détectable à la lumière blanche. Par exemple, ces techniques ont permis de visualiser des fibres de muscles squelettiques et de tissus conjonctif en utilisant la lectine WGA marquée avec un fluorophore (Kostrominova, 2011). De même, la lectine GNA a été utilisée pour l'étude du trafic au niveau des membranes (Morgan et al., 2013).

Les lectines peuvent aussi être des substituts aux anticorps dans les techniques d'immunotransfert. Par exemple, les lectines marquées initialement avec de la biotine peuvent permettre de révéler la présence d'une glycoprotéine donnée après analyse par électrophorèse (Morgan et al., 2013). Les lectines sont dans ce cas utilisées de la même manière que pourrait l'être un anticorps, avec l'avantage d'être plus faciles à préparer et facilement disponibles sous forme commerciale. Elles sont également utilisées dans des systèmes ELLA (Enzyme Linked Lectin Assay) qui est basé sur le même principe que l'ELISA (Enzyme-linked immunosorbent Assay) mais en utilisant des lectines au lieu des anticorps (McCoy Jr et al., 1983). La méthode directe permet de fixer directement les glycoprotéines sur les puits, puis d'ajouter la lectine marquée, sans utilisation d'anticorps. Ce n'est pas le cas de la méthode hybride où on fixe d'abord des anticorps au fond des puits pour capturer les glycoprotéines d'intérêt (*Figure I-14*).

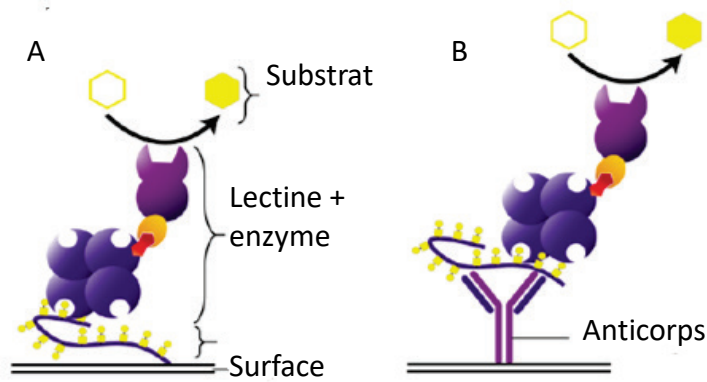


Figure I-14. Approches directes (A) et hybrides (B) des tests ELLA (Enzyme Linked Lectin Assay). D'après Hashim et al., 2017.

À plus grande échelle, la glycobiologie a besoin de plus en plus d'outils pour parvenir à comprendre la grande complexité du glycome au sein des cellules et des organismes. Comme décrit dans les sections précédentes, les glycanes ont des structures très variées et très complexes. Les lectines, de par leurs propriétés sont des outils de choix pour déchiffrer le « glycode ». Pour optimiser les analyses, les systèmes micro-puces à lectine ont été récemment développés. Le principe est d'immobiliser un panel de lectines sur une puce et d'y analyser des échantillons tels que des glycoprotéines, des cellules ou des bactéries. Cette technique a l'avantage d'être facile et rapide à mettre en œuvre et ne nécessite que très peu de matériel biologique. Les puces à lectines peuvent être utilisées dans diverses applications comme le suivi des marqueurs de cancers et autres maladies (Li et al., 2011), la distinction entre différents types cellulaires en fonction de leur glycome (Tateno et al., 2011) ou l'analyse rapide du profil glycosidique de glycoprotéines (Pilobello et al., 2005 ; **Figure I-15**).

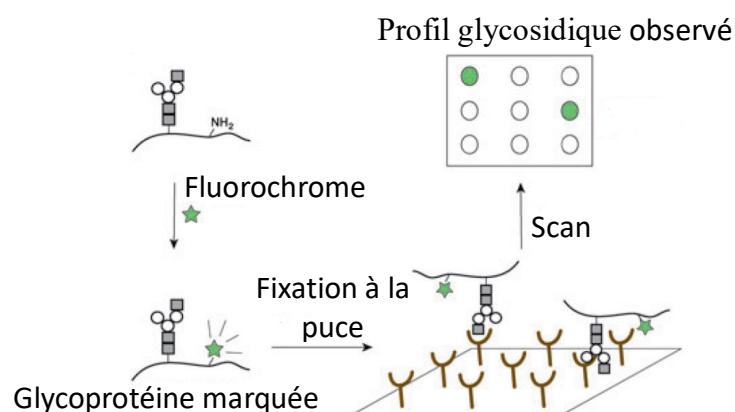


Figure I-15. Représentation schématique d'une expérience sur micro-puce à lectine. D'après Pilobello et al., 2005.

Les lectines peuvent aussi être utilisées comme biocapteurs, c'est-à-dire que les interactions entre les lectines et leurs ligands sont transformées sous forme de signal sans marquage préalable des lectines. Selon la technique, on peut suivre les variations de l'indice de réfraction lors de l'interaction (SPR : résonance plasmonique de surface), les modifications de la fréquence de résonance (QCM : microbalance à quartz), ou la réponse électrique (EIS : Spectroscopie d'impédance électrochimique) (La Belle et al., 2007 ; **Figure I-16**).

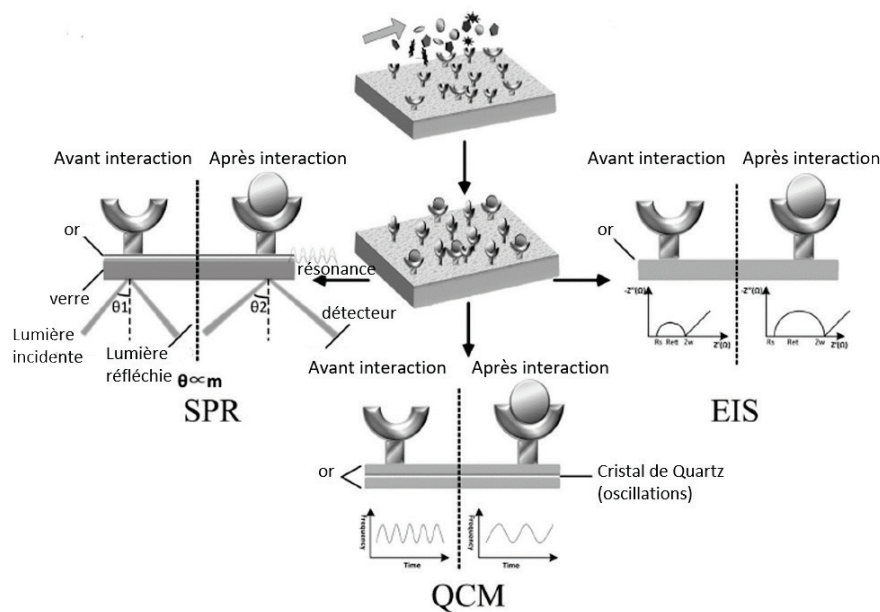


Figure I-16. Illustration des différents biocapteurs à base de lectine.
(Dan et al., 2016).

Ces différents outils d'analyse de la glycosylation peuvent s'appliquer au diagnostic et au suivi de cancers puisque ces derniers sont caractérisés par des modifications du glycome.

2.3.2 Diagnostic de cancers

La grande sélectivité et spécificité des lectines pour diverses structures de glycanes, font de ces protéines des outils à fort potentiel pour le diagnostic de cancers. Aujourd'hui, la majorité des marqueurs de cancers ciblés sont des glycoprotéines (Henry and Hayes, 2012). Pourtant, le suivi de la modification des glycanes des glycoprotéines ne se fait actuellement que sur l'alpha-fetoprotéine (AFP-L3) et CA15-3, des marqueurs de carcinome hépatocellulaire et de cancer du sein respectivement. Pour les autres glycoprotéines, ce sont les niveaux en protéines totales qui sont suivis, indépendamment de leur profil de glycosylation (Kuzmanov et al., 2013). Les lectines ont donc un rôle important à jouer dans le suivi des modifications de la glycosylation associée au cancer, que ce soit la sialylation, la fucosylation, les structures GlcNAc-branchées ou la surexpression de O-glycanes tronqués (Pinho and Reis, 2015).

Un exemple majeur de l'utilisation des lectines en tant qu'outils diagnostiques est la lectine LCA issue de *Lens culinaris*, et extraite de graines de lentilles. Cette lectine reconnaît l' α 1,6 fucose, et est utilisée pour le diagnostic du carcinome hépatocellulaire. Elle reconnaît spécifiquement l'alpha-fétoprotéine, et a été incorporée dans un kit commercial développé pour le diagnostic de l'HCC. La lectine LCA a été approuvée par la FDA (Food and Drug Administration ; agence américaine des produits alimentaires et médicamenteux) et est utilisée par de nombreux centres de traitement des cancers aux Etats-Unis (Pervin et al., 2015).

D'autres glycoprotéines marqueurs de cancers sont ciblées par des lectines à fort potentiel diagnostique. Ainsi, les deux glycoprotéines CA125 (cancer antigen 125) et HE4 (Human Epididymis protein 4) approuvées par la FDA, sont reconnues par un panel de lectines qui fixent les antigènes T et TF. Parmi elles, la lectine VVL (*Vicia villosa* lectine) permet de distinguer des tumeurs bénignes de tumeurs invasives dans le cas de cancers des ovaires épithéliaux (Chen et al., 2013). L'antigène spécifique de la prostate (PSA) est le biomarqueur le plus ciblé pour la détection de cancers de la prostate (Hayes and Barry, 2014). Il est reconnu par exemple par les lectines AAL (*Aleuria aurantia* lectin) et SNA (*Sambucus nigra* agglutinin) respectivement spécifiques du fucose et de l'acide sialique (Li et al., 2009). La Thyroglobuline (Tg) est une autre glycoprotéine approuvée par la FDA, cette fois en tant que biomarqueur dans les cancers de la thyroïde. Elle est reconnue par la lectine WFL (*Wisteria floribunda* lectine) qui fixe le GalNAc (Takeya et al., 2007).

L'utilisation des lectines permet aussi de découvrir de nouveaux biomarqueurs candidats. Cela a été le cas avec la lectine PHA-L qui a permis de découvrir 26 nouveaux biomarqueurs pour le diagnostic de cancers colorectaux (Kim et al., 2008). De plus, les lectines AAL et MAL-II sont des outils potentiels pour le diagnostic de ce type de cancers et le suivi de la maladie, puisqu'elles peuvent permettre de différencier des échantillons de tumeurs bénignes et malignes (Qiu et al., 2008). La distinction entre cancers du sein métastatiques et non métastatiques est également possible en utilisant les lectines AOL, GNA, RCA 120 et PHA sur des échantillons de sérum de patients (Fry et al., 2011). AOL (*Aspergillus oryzae* lectin) est spécifique du core fucose, GNA (*Galanthus nivalis* agglutinine) reconnaît préférentiellement les glycanes riches en mannose, RCA 120 (*Ricinus communis* agglutinine) fixe le galactose et PHA-E (*Phaseolus vulgaris* erythroagglutinine) reconnaît le GlcNAc. Ces quatre lectines montrent un profil de fixation différent selon que le sérum provient de patients atteints de cancers métastatiques ou non.

2.4 Les lectines recombinantes

2.4.1 Intérêt et mise en pratique

L'extraction des lectines de leur environnement naturel (plantes, champignons...) est difficile. Ce sont des manipulations qui prennent du temps et qui demandent d'avoir beaucoup de matière disponible au départ car les rendements peuvent être très faibles. De plus, les lectines peuvent être contaminées par d'autres biomolécules, dont des lectines que l'on ne souhaite pas purifier. Également, au sein de différents lots de plantes ou de champignons par exemple, on observe souvent une hétérogénéité de séquences des lectines et la présence d'isoformes qui peuvent avoir des propriétés différentes (Oliveira et al., 2013). C'est le cas par exemple des deux isoformes PHA-E et PHA-L (lectines issues de *Phaseolus vulgaris*), qui n'ont pas les mêmes propriétés d'agglutination des érythrocytes ((Raemaekers et al., 1999).

La production des lectines sous forme recombinante, c'est-à-dire en introduisant le gène de la protéine dans un organisme hétérologue (bactérie, cellule de mammifère, plantes...) est une stratégie qui permet d'obtenir de meilleurs rendements, une meilleure pureté, et de s'affranchir des problèmes liés aux isoformes. La production est également plus rapide et ne demande pas de matière première (Gemeiner et al., 2009), ce qui est un avantage important lorsque l'organisme produisant naturellement la lectine est difficile à trouver ou à cultiver ou présente des cycles saisonniers, comme c'est le cas pour la plupart des champignons.

Une méthode de choix pour la production de protéines sous forme recombinante est l'expression en bactéries, le plus souvent *Escherichia coli*. La croissance bactérienne étant très rapide, cela permet de réduire drastiquement les temps de production. *E. coli* est facile à cultiver et a un système génétique simple qui a permis d'en comprendre le fonctionnement et de l'optimiser pour la production de protéines à hauts rendements, généralement plusieurs milligrammes par litre (Oliveira et al., 2013), et à faibles coûts (Demain and Vaishnav, 2009). Ce système permet aussi de co-exprimer des protéines chaperonnes pour aider au repliement ou encore d'ajouter des étiquettes à l'extrémité N ou C-terminale de la protéine d'intérêt pour faciliter la purification. Les étiquettes N-terminales sont généralement clivables par des enzymes. Cependant l'expression en bactéries a aussi des désavantages car les bactéries ne permettent pas de réaliser de modifications post-traductionnelles telles que la glycosylation, qui sont parfois nécessaires à l'activité des lectines. C'est par exemple le cas pour la lectine PCL (issue de *Pleurotus cornucopiae*) qui a besoin d'être glycosylée pour être fonctionnelle (Sumisa et al., 2004).

Un système d'expression largement utilisé pour permettre notamment les modifications post-traductionnelles est la production en levures. En effet, ce sont des cellules eucaryotes qui ont donc des propriétés cellulaires plus proches de celles des organismes eucaryotes (plantes, champignons, animaux...) d'où sont généralement originaires les lectines. La lectine KM+ (issue de *Artocarpus integrifolia*) est un exemple de lectine qui a pu être produite avec succès dans la levure *Saccharomyces cerevisiae* avec un meilleur rendement que chez *E. coli* (daSilva et al., 2005). Un autre type de levure largement utilisé pour la production de protéines recombinantes est *Pichia pastoris*, qui permet de combiner les avantages des systèmes eucaryotes et procaryotes. Plus de deux mille protéines d'origines variées ont pu être exprimées dans ces levures (Cregg et al., 2000).

D'autres systèmes d'expression existent comme les cellules d'insectes, les cellules de mammifère ou les plantes, mais ils sont moins populaires que les bactéries et les levures (Oliveira et al., 2013). Cela est probablement dû à leurs nombreux désavantages comme la complexité et la durée des manipulations, des coûts plus élevés et de faibles rendements de l'ordre du microgramme par litre (Yin et al., 2007).

2.4.2 Particularité des lectines à pont disulfure

Les ponts disulfures sont formés par une réaction d'oxydation entre les groupements thiols (SH) de deux résidus cystéine. Au cours du repliement des protéines, ils peuvent être oxydés, réduits ou isomérisés (**Figure I-17**). Ces réactions sont réalisées par des oxydoréductases qui réalisent le transfert des électrons entre la protéine substrat et l'enzyme. Toutes les oxydoréductases ont un repliement de type thiorédoxine dont le site actif est constitué du motif CxxC avec x un acide aminé quelconque (Holmgren, 1972). Plusieurs facteurs vont déterminer si une oxydoréductase agit en tant qu'oxydase ou en tant que réductase : la stabilité relative de l'enzyme dans ses états réduits et oxydés, le compartiment dans lequel réside l'enzyme au sein de la cellule, ou encore la composition en acides aminés du site actif (Berkmen, 2012). À titre d'exemple, la thiorédoxine (Trx1) qui est une réductase cytoplasmique, peut agir comme une oxydase lorsqu'elle est exprimée dans le périplasma car elle y est isolée des réductases qui la maintiennent dans un état réduit dans le cytoplasme (Debarbieux and Beckwith, 1998).

Les ponts disulfures entre deux cystéines au sein de nombreuses protéines, sont cruciaux pour leur bon repliement et leur stabilité (Doig and Williams, 1991). Malheureusement, le cytoplasme natif des cellules d'*E. coli* ne permet pas la formation de ces ponts disulfures car il

comporte un grand nombre de réductases et d'agents réducteurs tel que le glutathion (Berkmen, 2012).

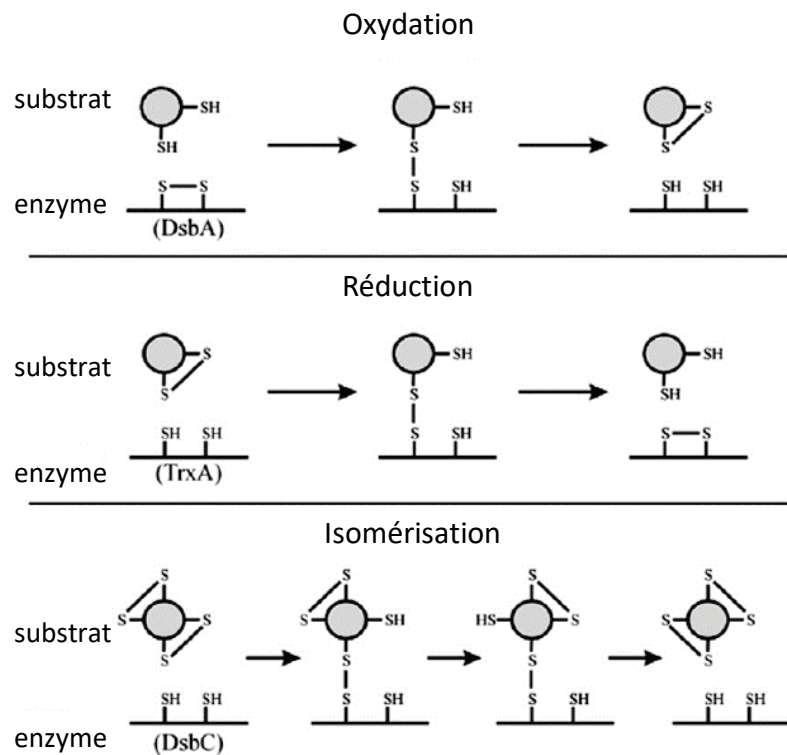


Figure I-17. Représentation schématique de transferts d'électrons entre une protéine substrat (ronds gris) et une oxydoréductase (lignes noires) permettant l'oxydation, la réduction ou l'isomérisation des ponts disulfures. D'après (Berkmen, 2012).

La première solution pour produire des lectines possédant des ponts disulfures dans *E. coli* est de les exprimer dans le périplasma. En effet, il s'agit d'un compartiment cellulaire oxydant dans lequel des protéines Dsb (« disulfide bond ») catalysent la formation et la correction des ponts disulfures (Kadokura et al., 2003). Cependant, cela nécessite de transloquer la protéine du cytoplasme vers le périplasma à travers la membrane interne. Cette étape délicate peut se faire via les deux voies de sécrétion Sec ou SRP (den Blaauwen and Driessen, 1996; Luirink and Sinning, 2004) en ajoutant à la séquence de la protéine une séquence signal donnée. La voie Sec est la plus commune et permet un transport post-traductionnel tandis que la voie SRP est un transport co-traductionnel. La protéine exportée dans le périplasma rencontre DsbA qui catalyse la formation de ses ponts disulfures. DsbA est maintenue dans son état actif par DsbB, une protéine de la membrane interne. Si la protéine a été mal oxydée par DsbA, elle est alors reconnue par DsbC et DsbG, deux isomérases maintenues dans un état réduit par DsbD (*Figure I-18* ; de Marco, 2009).

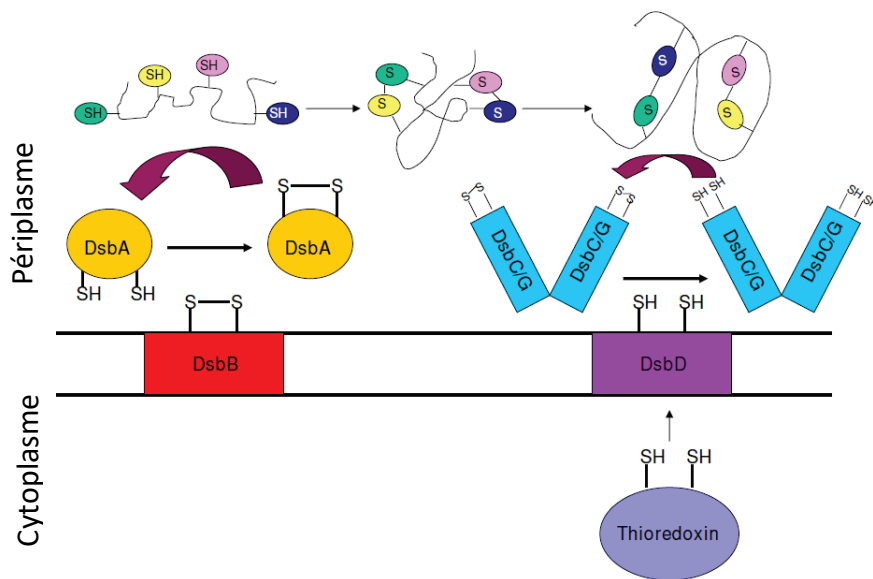


Figure I-18. Réactions d'oxydation et d'isomérisation des ponts disulfures dans le périplasm d'*E. coli*. D'après de Marco, 2009.

La deuxième solution est de produire les protéines possédant des ponts disulfures dans des cellules d'*E. coli* dont le cytoplasme a été modifié afin qu'il ne soit plus un environnement réducteur. Naturellement, cet environnement réducteur est maintenu par les voies glutaredoxine et thioredoxine (Prinz et al., 1997). La neutralisation de ces deux voies est létale pour les cellules, mais la viabilité peut être recouvrée par mutation de la peroxydase AhpC, de façon à maintenir une quantité minimale de glutathion essentielle à la croissance. L'environnement devient alors assez oxydant pour que s'accumulent des thioredoxines oxydées qui catalysent la formation des ponts disulfures (Ritz et al., 2001). Ces mutations sont réalisées dans les cellules Shuffle par exemple, qui co-expriment également DsbC pour permettre l'isomérisation des ponts disulfures mal formés (Tait and Straus, 2011).

3. Les lectines spécifiques de marqueurs cancéreux

3.1 Reconnaissance de la core-fucosylation

3.1.1 État de l'art

Comme cela a été précédemment décrit, le core-fucosylation est un très bon marqueur pour le diagnostic de cancers et le suivi de l'évolution de la maladie en fonction des traitements administrés. Le développement d'outils capables de reconnaître spécifiquement la core-fucosylation est donc essentiel pour pouvoir suivre les taux de core-fucosylation de glycoprotéines sériques par exemple. Cela permettrait aussi de développer des outils capables de réaliser le contrôle qualité des anticorps thérapeutiques dont la core-fucosylation diminue

l'efficacité. Plusieurs lectines spécifiques du fucose et capables de reconnaître la core-fucosylation ont déjà été étudiées. C'est le cas de AAL d'*Aleuria auratia*, AOL d'*Aspergillus oryzae*, LCA de *Lens culinaris* et PSA de *Pisum sativum* (Tateno et al., 2011).

AAL est une lectine fongique dimérique qui a pu être produite sous forme recombinante chez *E. coli* (Fukumori et al., 1990) et dont la structure tridimensionnelle est connue (Wimmerova et al., 2003). Chaque monomère a un repliement de type β -propeller à six lames (**Figure I-19.A**). AAL possède cinq sites d'interaction avec le fucose par monomère. Elle a une affinité de 1.6×10^{-4} M pour le fucose et reconnaît tous les oligosaccharides contenant du fucose indépendamment de sa liaison mais avec une préférence pour l' α 1,6 fucose. Malgré qu'elle ne soit pas strictement spécifique de la core-fucosylation, AAL a été beaucoup utilisée pour la reconnaissance de ce motif car il s'agissait de la seule lectine disponible (Wimmerova et al., 2003). AAL a une structure très proche de celle de AOL qui consiste également en un β -propeller à six lames, avec six sites d'interactions pour le fucose (Makyio et al., 2016 ; **Figure I-19.B**). AOL présente 26 % d'homologie de séquence avec AAL et une plus grande spécificité pour la core-fucosylation qu'AAL (Matsumura et al., 2007).

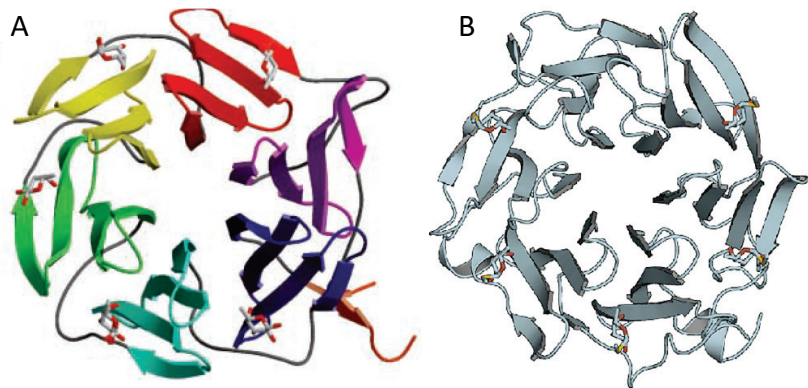


Figure I-19. Structure des lectines AAL (*Aleuria auratia*) et AOL (*Aspergillus oryzae*). Images issues respectivement de (Wimmerova et al., 2003) et (Makyio et al., 2016).

LCA et PSA sont deux protéines homologues qui reconnaissent des oligosaccharides α -mannosylés et notamment des N-glycanes mono- et bi-antennés avec une plus forte affinité en présence de core fucose. De plus, elles ne reconnaissent pas la core-fucosylation sur les N-glycanes tri- et tetra-antennés (Tateno et al., 2009). Les structures de LCA (Loris et al., 1994) et PSA (Prasthofer et al., 1989) ont été résolues par cristallographie. Elles forment des dimères et présentent un repliement de type β -sandwich. (**Figure I-20**).

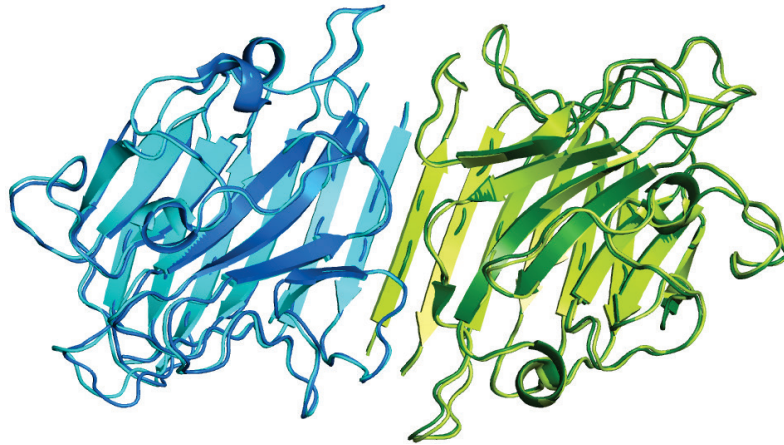


Figure I-20. Superposition des structures des lectines LCA et PSA.

Coloration des monomères en nuances de bleu et de vert (PSA foncé, LCA clair). PDB : 1LEN (LCA) et 2LTN (PSA).

Les lectines AAL, AOL, LCA et PSA ont donc des propriétés intéressantes pour la reconnaissance de la core-fucosylation mais elles ne sont pas spécifiques de ce motif. Pourtant en 2012, seule LCA était commercialisée pour la détection spécifique de la core-fucosylation, ce qui démontre l'importance de développer des outils plus spécifiques (Kobayashi et al., 2012). Seules cinq lectines ont pour le moment été décrites comme étant strictement spécifiques de la core-fucosylation α 1,6, il s'agit des lectines CSL issue de *Cephalosporium curvulum* (Inamdar et al., 2016), BTL issue de *Bryothamnion triquetrum* (do Nascimento et al., 2015), RSL issue de *Rhizopus stolonifer* (Oda et al., 2003), PhoSL issue de *Pholiota squarrosa* (Kobayashi et al., 2012) et SL2-1 issue de *Streptomyces rapamycinicus* (Vainauskas et al., 2016).

CSL est une lectine fongique qui a été purifiée à partir du champignon *Cephalosporium* qui est impliqué notamment dans les infections oculaires. CSL forme des tétramères dont les sous-unités font 14 kDa (Nagre et al., 2010). Sa spécificité pour la core-fucosylation des N-glycanes a été analysée sur une puce à sucre contenant 511 glycanes différents. Les résultats ont révélé sa stricte spécificité pour les N-glycanes core-fucosylés. Elle n'a pas d'affinité pour les O-glycanes fucosylés, et ne reconnaît pas les monosaccharides, fucose inclu. CSL présente un fort potentiel pour la reconnaissance de marqueurs de cancers, puisqu'elle reconnaît spécifiquement le core-fucose sur les cellules cancéreuses HepG2 (carcinome hépatocellulaire) et HT29 (cancer du colon) en cytométrie de flux. Elle a également un fort potentiel diagnostique puisqu'elle reconnaît fortement les tissus cancéreux et métastatiques et très peu les tissus sains du colon (Inamdar et al., 2016). Cependant, cette lectine n'est pas

disponible sous forme recombinante, ce qui limite fortement son utilisation en tant qu'outil diagnostique et sa séquence primaire est indisponible à ce jour.

BTL est une lectine d'algue monomérique de 10 kDa. La détermination de la séquence de cette lectine a révélé qu'elle n'avait aucune similarité avec des lectines connues, et qu'elle est donc la première représentante d'une nouvelle famille de lectines (Calvete et al., 2000). Elle est strictement spécifique de la core-fucosylation mais elle a une préférence pour les N-glycanes bi- et tétra-antennés (analyse sur puce à sucres). Elle a une affinité de 12 μM pour le core-fucose qui a été déterminée par titration calorimétrique isotherme (ITC). Elle a pu être produite sous forme recombinante dans *E. coli* mais avec un faible rendement de 1 mg par litre de culture. De plus, la lectine recombinante est moins active que sa forme native et est instable dans le temps. Malgré les efforts, aucune structure tridimensionnelle n'a pu être résolue pour cette lectine (do Nascimento et al., 2015).

RSL est une petite lectine fongique de seulement 4.5 kDa qui forme des hexamères et qui reconnaît spécifiquement la core fucosylation. La spécificité de cette lectine a été analysée par des tests d'inhibition d'hémagglutination et par électrophorèse capillaire d'affinité. Elle présente cependant de faibles affinités pour le L-fucose puisque le sucre seul est capable d'inhiber l'activité hémagglutinante de cette lectine (Oda et al., 2003). La séquence primaire de RSL a été déterminée et montre 85 % d'homologie avec la séquence de PhoSL décrite ci-dessous (Kobayashi et al., 2012) mais elle n'est pas disponible sous forme recombinante.

SL2-1 est une lectine bactérienne monomérique strictement spécifique de la core-fucosylation, découverte par similarité de séquence avec RSL et PhoSL. Elle présente au sein de sa séquence trois répétitions d'une séquence d'environ 40 acides aminés, homologue à celle de PhoSL. Elle a été exprimée sous forme recombinante dans *E. coli* avec succès et sa spécificité a été vérifiée sur puce à sucre également. Son affinité pour le core-chitobiose est de $1.59 \cdot 10^{-6}$ M et elle semble avoir deux sites de liaison par monomère (mesuré par ITC). Le chitobiose core-fucosylé est le plus petit épitope reconnu avec une forte affinité par SL2-1. SL2-1 est capable de reconnaître la core-fucosylation des N-glycanes des IgG et de reconnaître spécifiquement les N-glycanes core-fucosylés dans un mélange complexe (Vainauskas et al., 2016). Il s'agit donc d'un outil prometteur pour la reconnaissance stricte de la core-fucosylation.

Enfin, PhoSL est une petite lectine fongique de 4,5 kDa qui forme des trimères et qui présente 85 % d'homologie de séquence avec RSL. C'est une protéine très stable et dont la

stricte spécificité pour les N-glycanes core-fucosylés a été vérifiée par chromatographie d'affinité frontale (PhoSL est immobilisée sur une résine sur laquelle sont déposés différents sucres à tester). Cette technique a aussi permis de mesurer des affinités de l'ordre du micromolaire pour les N-glycanes core-fucosylés. PhoSL a un fort potentiel en tant qu'outil de marquage de la core-fucosylation : elle reconnaît spécifiquement les glycoprotéines core-fucosylées et notamment des marqueurs de cancers tels que l'alpha-fétoprotéine (AFP) associée aux cancers du foie. Elle est aussi capable de se fixer sur des coupes de cancer du côlon avec des différences de marquage entre des coupes métastatiques et primaires (Kobayashi et al., 2012). La mise au point d'un test PhoSL-ELISA est également très prometteur pour la reconnaissance de l'haptoglobine et donc le diagnostic de cancers du pancréas (Kusama et al., 2017). De plus, il a été montré que la création d'un hybride PhoSL-anthraquinone permettait de dégrader spécifiquement l'alpha-fétoprotéine et d'inhiber la croissance de cellules cancéreuses HuH-7 dérivées d'HCC (Okuyama et al., 2016). Cependant, cette lectine n'avait jamais été produite sous forme recombinante, ce qui est encore une fois un frein à son utilisation dans de multiples applications.

Au départ de ce travail, aucune structure tridimensionnelle n'était disponible pour aucune de ces lectines strictement spécifiques de la core-fucosylation, et la compréhension de cette spécificité à l'échelle moléculaire était donc impossible.

3.1.2 PhoSL et PhoNL

L'étude des protéines strictement spécifique de la core-fucosylation reste à approfondir, notamment leur production sous forme recombinante et leur étude structurale. PhoSL présente un grand intérêt, de par sa forte spécificité et son fort potentiel applicatif. Sa séquence a donc été utilisée pour rechercher des lectines homologues qui pourraient avoir des propriétés similaires et ainsi augmenter le nombre d'outils potentiels pour la reconnaissance de la core-fucosylation. Cela a permis de découvrir une deuxième lectine fongique du genre *Pholiota* et issue du champignon *Pholiota nameko* appelée PhoNL, dont seule la séquence en acides aminés est connue. L'alignement de séquence montre de fortes homologies entre PhoSL, PhoNL, RSL et SL2-1, des lectines qui appartiennent vraisemblablement à la même famille (**Figure I-21**). PhoSL et RSL sont des petites protéines de seulement 4.5 kDa tandis que PhoNL et SL2-1 font environ 17 kDa et présentent au sein de leur séquence trois répétitions homologues à la séquence de PhoSL. La trimérisation de PhoSL laisse ainsi penser qu'elle pourrait avoir une organisation tridimensionnelle homologue à celle de PhoNL, renforçant l'hypothèse d'avoir des propriétés similaires entre les deux lectines.

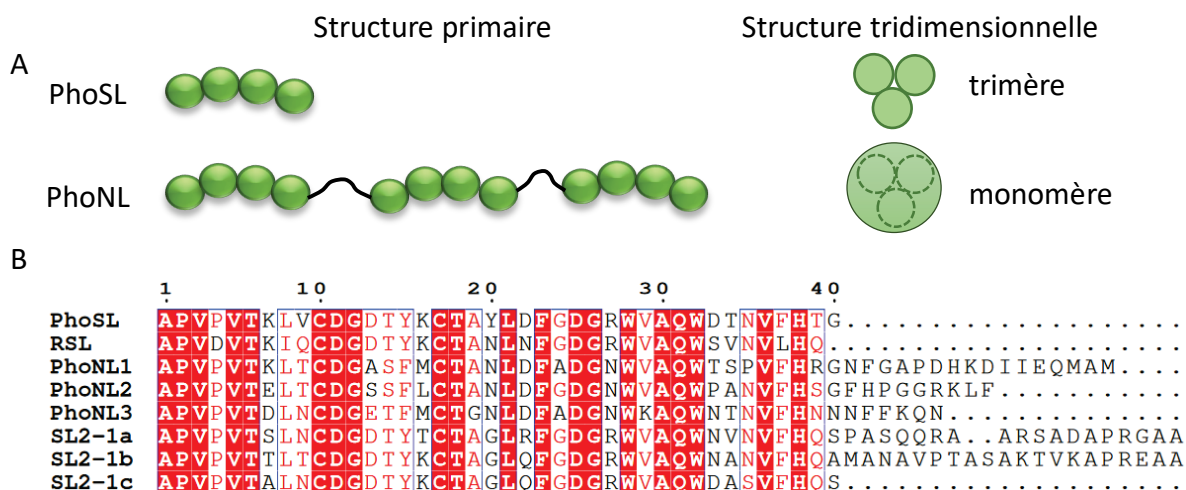


Figure I-21. Schématisation des différences de structure primaire et tridimensionnelle entre PhoNL et PhoSL (A) et alignement de séquence avec RSL et SL2-1 (B).

A. Le monomère de PhoNL présente trois répétitions de la séquence de PhoSL reliées entre elles par de courtes séquences de 7 à 16 acides aminés. PhoSL forme des trimères en solution et pourrait donc avoir une structure tridimensionnelle similaire à celle de PhoNL. B. Alignement de séquence. Les trois répétitions de PhoNL sont nommées PhoNL1, 2 et 3 et les trois répétitions de SL2-1 sont nommées SL2-1a, b et c.

Le fait que les trois répétitions de PhoNL soient présentes au sein d'un même monomère pourrait aussi faciliter la production recombinante de cette lectine comparé à PhoSL. De plus, la comparaison des propriétés et de la séquence de PhoNL avec les autres lectines PhoSL, RSL et SL2-1 contribuerait à élargir les connaissances sur cette famille de lectine.

3.2 Reconnaissance de l'antigène TF

3.2.1 État de l'art

L'antigène TF (Gal β 1-3GalNAc α 1-Ser/Thr) est un autre biomarqueur d'intérêt dans l'étude et le suivi des cancers notamment. Les lectines les plus pertinentes pour la reconnaissance de l'antigène TF sont aujourd'hui toutes issues de champignons. On peut en citer principalement quatre qui sont XCL de *Xercomus chrysenteron* (Birck et al., 2004), ABL d'*Agaricus bisporus* (Carrizo et al., 2005), SRL de *Sclerotium rolfsii* (Leonidas et al., 2007) et BEL de *Boletus edulis* (Bovi et al., 2011). Ces quatre lectines homologues appartiennent à la même famille et présentent un repliement de type actinoporine. Elles sont aussi caractérisées par la présence de deux sites de fixation des sucres de spécificité différente : un pour le GalNAc et un pour le GlcNAc.

XCL fut la première lectine de cette famille pour laquelle la structure 3D a été résolue. Elle est devenue la représentante d'une nouvelle famille de lectines fongiques d'environ 140

acides aminées présentant une spécificité forte pour le GalNAc. Elle a pu être produite sous forme recombinante dans *E. coli* et sa structure a permis de mettre en évidence une nouvelle architecture pour les lectines qui est curieusement très similaire au repliement des actinoporines ou cytolysines (des toxines produites par les anémones de mer), ce qui indique une relation évolutive entre ces protéines. Ce repliement est caractérisé par la présence d'un β -sandwich dont les deux feuillets β sont liés entre eux par un motif hélice-tour-hélice. XCL forme des tétramères et a une forte affinité de 1 μ M pour l'antigène TF mais elle reconnaît aussi le galactose, le GalNAc et le lactose avec une faible affinité (Damian et al., 2005). Ces spécificités lui confèrent des propriétés insecticides, par exemple envers *Drosophila melanogaster* et *Acrythosiphon pisum* (Trigueros et al., 2003).

ABL est une lectine tétramérique de 64 kDa pour laquelle la présence d'un deuxième site de fixation spécifique du GlcNAc a été mise en évidence (**Figure I-22**) (Irazoqui et al., 1999; Carrizo et al., 2005). Des analyses par chromatographie d'affinité frontale ont permis de mesurer une affinité de l'ordre du micromolaire pour l'antigène TF et une affinité d'environ 20 μ M pour le GlcNAc des N-glycanes. Elle est également capable de reconnaître les structures noyau de type 2 des O-glycanes (Gal β 1-3(GlcNAc β 1-6)GalNAc α) avec une affinité du même ordre de grandeur que pour le GlcNAc (Nakamura-Tsuruta et al., 2006). Sa double affinité pour les O-glycanes (GalNAc) et les N-glycanes (GlcNAc) limite son utilisation pour la reconnaissance spécifique de l'antigène TF. De plus, cette lectine n'est pas disponible sous forme recombinante mais elle a cependant d'excellentes propriétés pour l'inhibition de la prolifération de cellules cancéreuses épithéliales, sans cytotoxicité apparente pour les cellules normales (Yu et al., 1993).

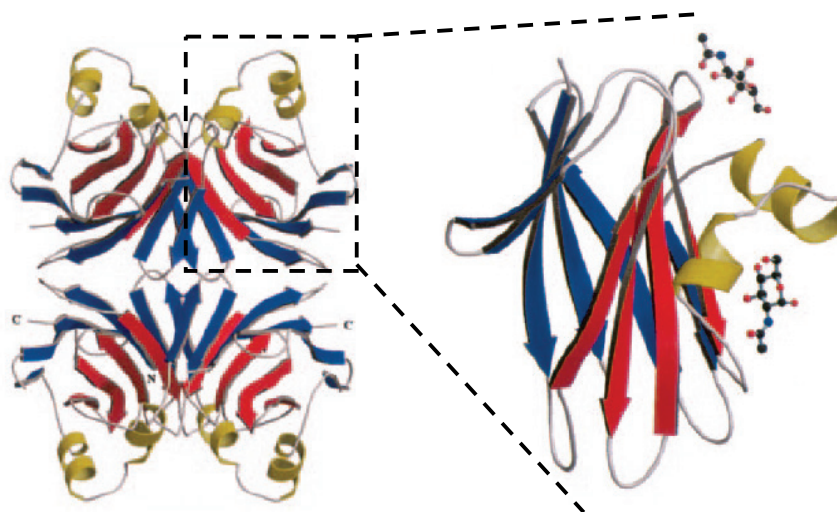


Figure I-22. Structure du tétramère de ABL avec un zoom sur un des monomères.
Image reprise de Carrizo et al., 2005. PDB : 1Y2W.

SRL a une taille de 17 kDa et forme principalement des dimères en solution. Elle pourrait avoir la capacité de former également des tétramères d'après sa structure cristallographique (Leonidas et al., 2007). SRL a déjà été produite sous forme recombinante dans *E. coli* (Vanti et al., 2016). Elle reconnaît l'antigène TF avec une forte affinité, mais reconnaîtrait également l'antigène Tn et faiblement le GalNAc (Wu et al., 2001). Une analyse sur puce à sucre de 2011 a conclu cependant que SRL ne reconnaît pas le GlcNAc sur les N-glycanes, ni le GalNAc seul, en considérant que seuls les signaux supérieurs à 50% du signal maximal sont significatifs. Cela indique la très faible affinité de SRL pour ces deux motifs. SRL est aussi capable de reconnaître fortement les noyaux de type 2 des O-glycanes et fixe plus spécifiquement l'antigène TF α qui est l'anomère le plus abondant sur les glycoprotéines (Chachadi et al., 2011). Deux variants de SRL, SSR1 et SSR2 ont par la suite été produits afin d'affiner la spécificité de cette lectine. Ces deux variants reconnaissent l'antigène TF de la même manière que SRL mais SSR1 reconnaît en plus l'antigène Tn ainsi que sa forme sialylée (Peppas et al., 2015). Enfin, il a été démontré que SRL est capable d'inhiber la croissance de cellules cancéreuses du côlon, et pourrait donc permettre de développer de nouveaux agents thérapeutiques contre certains cancers (Inamdar et al., 2012).

BEL est aussi homotétramérique et possède 82 % d'identité de séquence avec XCL, 57 % avec ABL et 58 % avec SRL. Elle présente des propriétés antiprolifératives pour des cellules de carcinomes tel que HT29 (côlon), HepG2 (foie) et MCF7 (sein) grâce à la reconnaissance de l'antigène TF (Bovi et al., 2011). BEL n'a pas été produite sous forme recombinante et peu de publications décrivent davantage les spécificités de cette lectine, et ses applications possibles en diagnostic et en thérapie.

3.2.2 TrfbL2

Au cours d'un travail précédent, la lectine fongique TrfbL2 issue de *Terfezia boudieri* (truffe du Sahara Algérien) a été extraite, purifiée, caractérisée au CERMAV. C'est la première lectine, avec TrfbL1, une lectine fucose spécifique, à avoir été caractérisée dans ce champignon.

L'étude de TrfbL2 a montré qu'il s'agissait d'une protéine très stable avec une température de fusion autour de 80°C qui ne reconnaissait pas les sucres simples dont le GalNAc et le GlcNAc. Une analyse plus large de sa spécificité sur une puce à sucre a montré que TrfbL2 reconnaissait les GlcNAc terminaux des N-glycanes ainsi que les glycanes possédant le motif Gal β 1-3GalNAc α qui correspond à l'antigène TF. L'affinité pour l'antigène TF mesurée par

SPR (résonance plasmonique de surface) était de 3 μM tandis que celle pour le GlcNAc était de 0.6 mM. L'analyse de sa séquence N-terminale a mis en évidence son homologie avec la famille de lectines de type actinoporines. Sa structure a été résolue par remplacement moléculaire avec XCL comme modèle à résolution atomique, ce qui a permis d'obtenir la séquence primaire de la protéine à partir de la densité électronique (**Figure I-23**). Les structures en complexe avec l'antigène TF ou du GlcNAc et le disaccharide GlcNAc β 1-3Gal ont aussi été résolues.

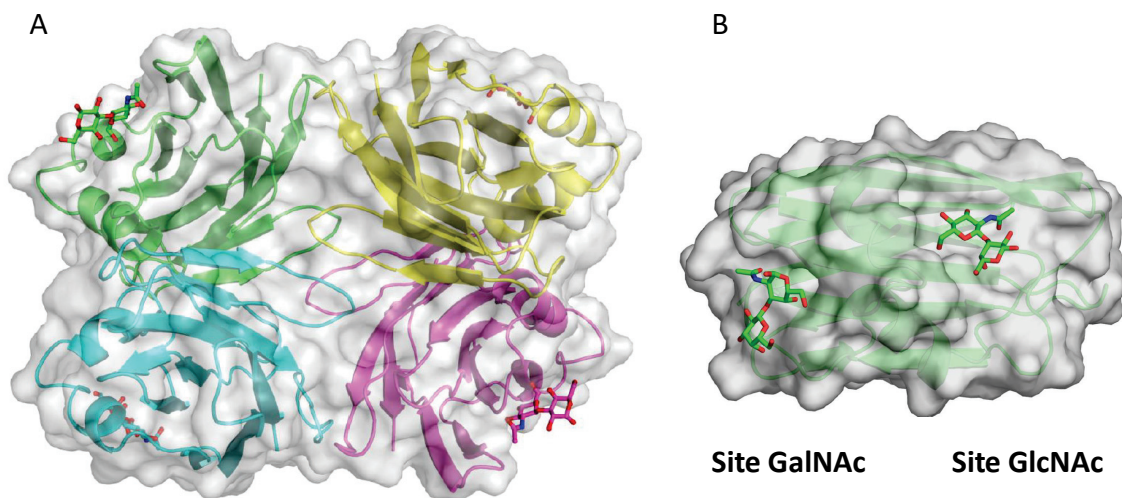


Figure I-23. Structure du tétramère de TrfbL2 (A) et d'un monomère avec les sucres en interaction sur les sites à GlcNAc et à GalNAc (B).

Les sucres en interaction sont l'antigène TF sur le site GalNAc et le disaccharide GlcNAc β 1-3Gal sur le site GlcNAc.

La détermination de la séquence de TrfbL2 a permis par la suite d'isoler le gène de cette protéine afin de la produire sous forme recombinante dans *E. coli*. Une des limites de l'utilisation des lectines de cette famille pour la reconnaissance de l'antigène TF est la double spécificité pour le GalNAc et pour le GlcNAc. Bien que l'affinité pour le GlcNAc soit 200 fois plus faible que pour l'antigène TF, elle reste significative et pourrait entraver la reconnaissance spécifique de l'antigène TF. La production de TrfbL2 sous forme recombinante pourrait permettre de réaliser des mutants encore plus spécifiques et augmenter son potentiel en biomarqueur.

II.

OBJECTIFS DE LA

THÈSE

Les aberrations de la glycosylation sont caractéristiques de nombreux cancers, car ils sont retrouvés à des taux anormaux en surface de nombreuses cellules cancéreuses. Le développement d'outils capables de reconnaître spécifiquement ce type d'aberration est donc devenu d'un intérêt croissant en recherche et diagnostic des cancers. Ils peuvent aussi être utiles aux domaines des biotechnologies pour l'analyse d'échantillons complexes ou de la glycosylation de glycoprotéines, notamment suite à l'essor de la glycomique. Parmi les marqueurs de cancers existants, notre objectif a été focalisé sur la core fucosylation et l'antigène TF. L'objectif est de développer des lectines recombinantes faciles à produire, strictement spécifiques de ces marqueurs et qui pourront compléter et améliorer les outils de marquage déjà existants qui ne sont pas toujours optimaux.

Au départ de ce travail de thèse, nous avons à disposition deux lectines spécifiques de la core-fucosylation, PhoSL et PhoNL, et une lectine spécifique de l'antigène TF, TrfbL2. J'ai produit ces lectines au laboratoire sous forme recombinante dans *E. coli* lors de mon stage de Master II. Des mutants de rTrfbL2 avaient été clonés dans le but d'invalider le site de reconnaissance du GlcNAc lors de mon stage de Master I. Les objectifs pour rPhoSL et rPhoNL sont :

- Affiner l'analyse des lectines d'un point de vue biochimique : stabilité, capacité d'héماغglutination, état d'oligomérisation.
- Analyser l'interaction des lectines avec leur ligand : affinités et spécificités pour des sucres core-fucosylés libres ou liés à des protéines.
- Résoudre les structures cristallographiques en absence et en présence de ligands pour comprendre les déterminants essentiels à l'interaction.
- Éventuellement utiliser les informations biochimiques et structurales pour optimiser ces lectines en réalisant des mutations ciblées.

Les objectifs pour rTrfbL2 sont :

- Produire les mutants et développer les protocoles de purification adaptés.
- Analyser, en comparaison avec la lectine sauvage, leurs caractéristiques biochimiques et l'interaction avec leurs ligands (évaluer leur spécificité pour l'antigène TF et l'invalidité du site de reconnaissance pour le GlcNAc).
- Résoudre les structures cristallographiques pour vérifier l'impact des mutations et les optimiser.

Enfin, l'objectif final est d'évaluer le potentiel de ces lectines dans le marquage de cellules et de tissus biologiques par des analyses de cytométrie en flux et d'histochimie.

III.
MATÉRIEL ET
MÉTHODES

1. Clonage et mutagenèse dirigée

rPhoSL et *rPhoNL*

Les gènes *rPhoSL* et *rPhoNL* ont été synthétisés avec une séquence optimisée pour une expression chez *Escherichia coli* (*E. coli*) et insérés dans le plasmide pUC57-Kan (Genscript USA). Le gène de *rPhoNL* contenait une séquence codant pour un peptide signal eucaryote qui a été ensuite éliminé par PCR (Réaction de Polymérisation en Chaîne). En raison de la prédiction de ponts disulfures, à la fois par le prédicteur Disulfind (Ceroni et al., 2006) et par l'analyse des protéines homologues, la production périplasmique de ces protéines a été retenue. Pour cela, les gènes *rPhoSL* et *rPhoNL* ont été sous-clonés dans le vecteur pET39b (Novagen) pour lequel le site de clivage de l'entérokinase a été remplacé par le site de clivage de la TEV (Tobacco Etch virus) protéase. Ce vecteur (pET39b-TEV) possède le gène de la protéine-fusion DsbA (**Figure III-1**) qui présente un signal de localisation dans le périplasma et y catalyse la formation des ponts disulfures. Suite à une double digestion avec les enzymes de restriction XhoI et NcoI, les gènes ont été insérés par ligation dans le plasmide en utilisant le Mighty Ligation Mix (Takara). Des PCRs sur colonie ont alors été réalisées, suivies d'une amplification dans la souche Top10 d'*E. coli* (Invitrogen) et de minipreps avec le kit «QIAprep Spin Miniprep» en suivant les instructions du fournisseur (QIAGEN) avant vérification par séquençage (Eurofins Genomics).

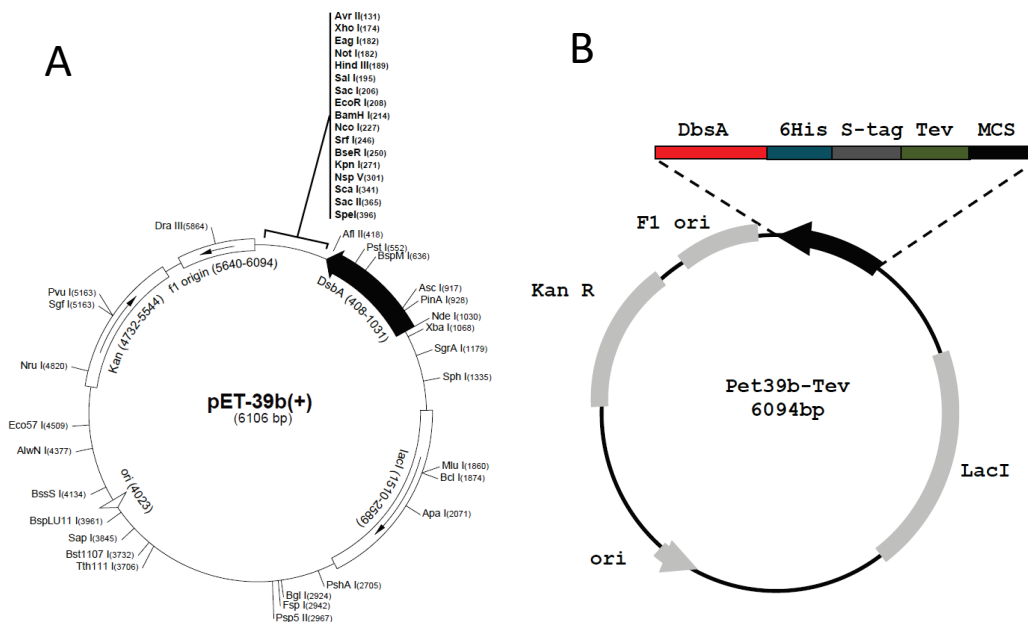


Figure III-1. A. Carte du vecteur pET39b (Invitrogen) ; B. caractéristiques du vecteur modifié pET39b-TEV.

Le mutant rPhoSLm a été conçu par mutagenèse dirigée à partir du clone de rPhoSL présent dans le plasmide Pet39b-TEV, afin de supprimer les premiers acides aminés alanine et méthionine de la séquence finale de rPhoSL. Cette mutation a été réalisée par PCR en suivant le protocole de mutagenèse QuickChange site directed mutagenesis (Stratagen) et en utilisant la polymérase PrimeStar Max (Ozyme). Les amorces nécessaires à la mutation ont été synthétisées par Eurofins Genomics.

Le mutant rPhoNLM est le résultat de plusieurs modifications qui sont détaillées *Chapitre IV*. Ce gène a été cloné dans le plasmide pET39b-TEV comme pour rPhoSL, puis les premiers acides aminés alanine et méthionine ont été supprimés selon la même méthode que celle utilisée pour la mutation rPhoSLm.

Le mutant rPhoNLMSS a été conçu à partir de la séquence du mutant rPhoNLM. Il est le résultat de la mutation de la serine 70 en cystéine selon le même protocole que pour rPhoSLm.

TrfbL2

Le gène de rTrfbL2 sauvage (rTrfbL2-WT) a été cloné dans le vecteur pET29a (Novagen) entre les sites NcoI et XhoI (*Figure III-2*). Il comporte un gène de résistance à la kanamycine et ne possède pas d'étiquette.

Toutes les amorces utilisées sont détaillées dans l'*Annexe*.

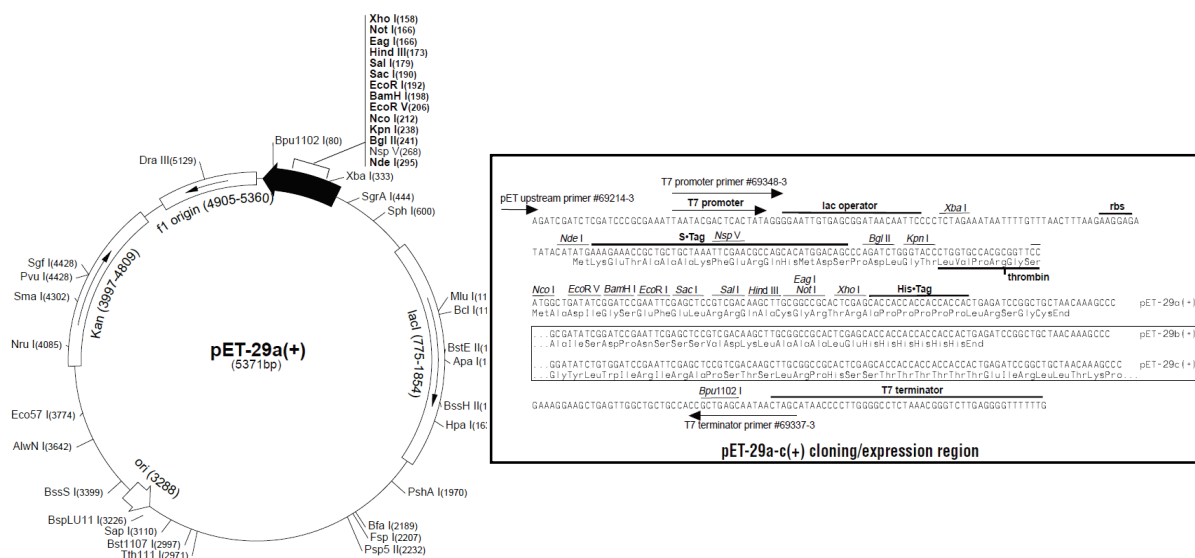


Figure III-2. Carte du vecteur pET29a.

Les mutants rTrfbL2-Y111A, Y111Q, Y111QD77A, D77A et R100G ont été conçus par mutagenèse dirigée à partir du plasmide rTrfbL2-pET29a. Ces mutations ont été réalisées par PCR en suivant le protocole de mutagenèse dirigée cité ci-dessus. Par la suite, les mutants rTrfbL2-D77A et rTrfbL2-Y111QD77A ont été clonés dans différents vecteurs d'expression (Pet-TEV, pET32-TEV, pProEx-Htb) afin de faciliter leur future purification. Ces vecteurs, contrairement à pET29a, permettent l'ajout de différentes étiquettes en N-terminal de la protéine. Leurs caractéristiques sont décrites *Figure III-3* et *Tableau III-1*.

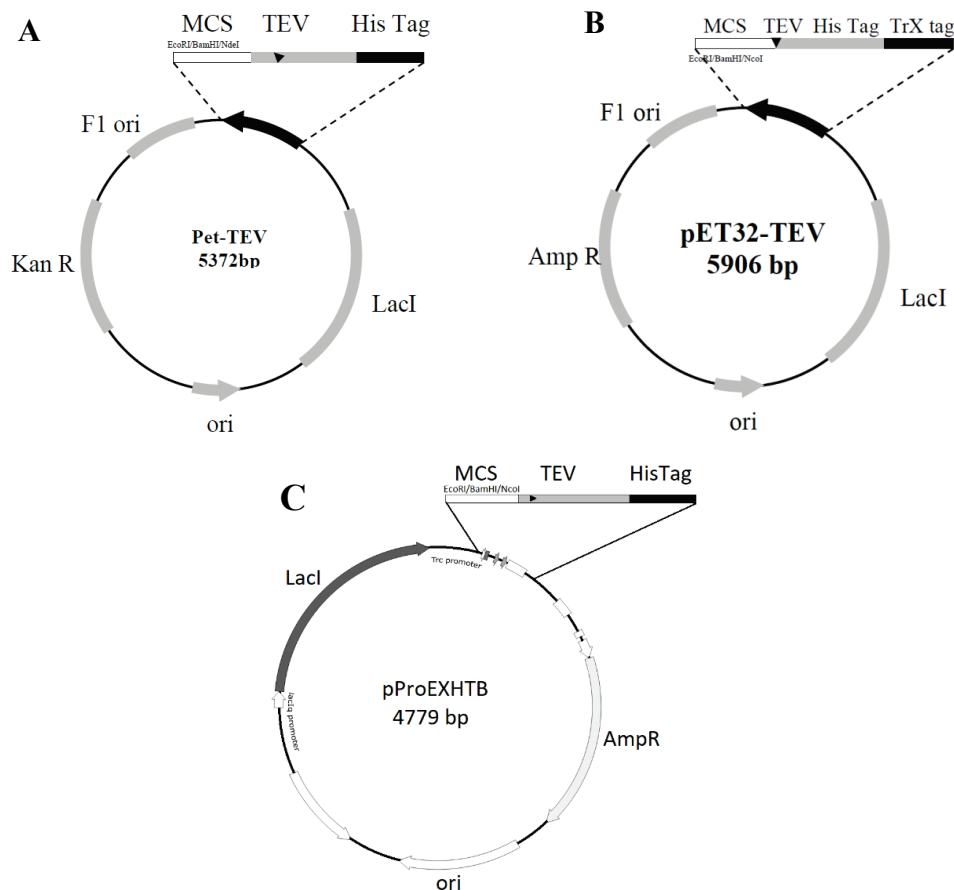


Figure III-3. Plasmides utilisés pour l'expression des mutants de rTrfbL2. Clonage entre les sites NcoI et XhoI du site de clonage multiple.

Vecteur	Promoteur	Etiquette	Gène de résistance
pProEx Htb	<i>trc</i>	histidine	ampicilline
pET-TEV	T7	histidine	kanamycine
pET32-TEV	T7	histidine + thioredoxine	ampicilline

Tableau III-1. Caractéristiques des vecteurs pProEx Htb, pET-TEV et pET32-TEV utilisés pour le clonage des mutants rTrfbL2-D77A et rTrfbL2-Y111QD77A.

Les promoteurs T7 et *trc* fonctionnent sur le même principe que le promoteur *lac* dont ils sont dérivés. Ils sont inductibles à l'isopropyl-β-D-thiogalactopyranoside (IPTG) qui est un analogue non hydrolysable du lactose (Terpe, 2006). Le promoteur T7 est plus fort que le promoteur *trc*, c'est-à-dire qu'il a une fréquence d'initiation de la transcription plus grande. De plus, il permet d'éviter les fuites d'expression si l'on inhibe l'ARN polymérase T7 par le lysozyme T7. Cependant, l'utilisation du promoteur *trc* permet parfois d'avoir une meilleure production avec une plus grande proportion de protéines solubles (Tegel et al., 2011).

2. Production et purification

2.1 Production et extraction

L'ensemble des cellules utilisées pour cette étude ainsi que leurs caractéristiques sont décrites dans le **Tableau III-2**. Toutes ces cellules sont dérivées de souches d'*E. coli*. Concernant les souches BL21Star(DE3) et Rosetta(DE3), la désignation « DE3 » signifie qu'elles contiennent le lysogène λDE3 qui porte le gène de l'ARN polymérase T7 sous le contrôle du promoteur lacUV5. Ces cellules portent aussi une mutation de la protéase OmpT, permettant de réduire la dégradation des protéines hétérologues lors de l'expression.

Cellules	Souche parentale	Génotype	Description
BL21 Star (DE3)	B834	$F^- ompT hsdS_B (r_B^- m_B^-) gal dcm rne131$ (DE3)	La mutation du gène codant pour la RNase E (<i>rne131</i>) permet une plus grande stabilité des ARNm.
Rosetta (DE3) pLysS	BL21	$F^- ompT hsdS_B (r_B^- m_B^-) gal dcm$ (DE3) <i>pLysSRARE</i> (<i>Cam^R</i>)	Le plasmide pLysS résistant au chloramphénicol porte six gènes d'ARNt codant pour des codons rares. Il est aussi porteur du lysosyme T7 qui supprime l'expression basale de l'ARN polymérase T7
Shuffle T7	K12	$F^+ lac, pro, lacI^q / \Delta(ara-leu)7697 araD139 fhuA2 lacZ::T7 gene1 \Delta(phoA)PvuII phoR ahpC^* galE (or U) galK \lambda att::pNEB3-r1-cDsbC$ (<i>Spec^R</i> , <i>lacI^q</i>) $\Delta trxB rpsL150(Str^R) \Delta gor \Delta(malF)3$	Les mutations des gènes <i>trxB</i> et <i>gor</i> créent un environnement cytoplasmique moins réducteur. La surexpression de DsbC (isomérase des ponts disulfures) assure un bon repliement des protéines

Tableau III-2. Description des cellules utilisées dans cette étude.

2.1.1 rPhoSL, rPhoNL, rPhoSLm, rPhoNLm

Afin de produire rPhoSL, rPhoSLm, rPhoNL et rPhoNLm, les plasmides d'intérêt ont été transformés par choc thermique à 42°C dans la souche d'expression BL21Star(DE3) d'*E. coli*

pour une expression périplasmique. Les précultures et cultures ont été réalisées en milieu LB en présence de l'antibiotique adéquat (kanamycine 30 $\mu\text{g}\cdot\text{ml}^{-1}$) à 37°C. Lorsque les cultures ont atteint une densité optique à 600nm (DO_{600}) égale à 0,4, elles ont été placées à 16°C. La production des protéines a été induite avec 0,1 mM d'IPTG lorsque la densité optique à 600 nm atteignait 0,7-0,9. Après une nuit à 16°C, les cultures ont été centrifugées à 5000 g pendant 10 minutes puis les culots ont été directement traités sans étape de congélation.

L'extraction périplasmique a ensuite été réalisée par resuspension des culots dans 30 mL de tampon composé de 30 mM Tris-HCl, 20% sucrose et 1 mM d'EDTA (acide éthylènediaminetétraacétique) pH 8. Les cellules ont été agitées à température ambiante pendant 10 min puis centrifugées 10 min à 10000g. Après resuspension dans 30 mL de 5 mM MgSO_4 froid, les cellules ont été agitées doucement 10 min dans la glace, puis centrifugées à nouveau 10 min à 10000 g. Le surnageant, constituant l'extrait soluble, a été filtré sur un filtre seringue de 0.45 μm puis conservé avant purification.

2.1.2 rPhoNLmSS

La production de rPhoNLmSS a été réalisée selon deux méthodes différentes. Pour la première, le plasmide pET39b-TEV a été transformé dans les cellules BL21Star(DE3). Puis les cultures et extractions périplasmiques ont été réalisées selon le protocole décrit ci-dessus pour rPhoNLm. Pour la seconde, le plasmide pET39b-TEV a été transformé dans les cellules Shuffle. Les précultures et cultures ont été réalisées en milieu LB en présence de kanamycine 30 $\mu\text{g}\cdot\text{ml}^{-1}$ à 30°C. Lorsque les cultures ont atteint une DO_{600} de 0,4, elles ont été placées à 16°C. La production des protéines a été induite avec 0,1 mM d'IPTG lorsque la DO_{600} était de 0,7-0,9. Après une nuit à 16°C, les cultures ont été centrifugées à 5000g pendant 10 min et le culot a été pesé. Afin de réaliser une extraction complète (cytoplasmique + périplasmique), les culots ont été resuspendus dans du tampon composé de 50 mM Tris pH 8,5 + 300 mM NaCl à raison de 5 mL de tampon par gramme de cellules obtenu. Les cellules ont ensuite été cassées avec un broyeur cellulaire une fois à une pression de 1.9 KPa (Constant System Ltd, UK) et centrifugées à froid à 24000 g pendant 30 min. Le surnageant a été filtré sur un filtre seringue de 0.45 μm et constituait l'extrait soluble.

2.1.3 rTrfbL2 et ses mutants

Pour produire rTrfbL2 sauvage ou mutée, le plasmide d'intérêt a été transformé dans les cellules RosettaLysS(DE3) ou BL21Star(DE3) par choc thermique à 42°C. Les précultures (30 ml) et cultures (1L) ont été réalisées en milieu LB en présence de kanamycine (30 $\mu\text{g}\cdot\text{mL}^{-1}$)

¹) et de chloramphénicol (34 $\mu\text{g}\cdot\text{mL}^{-1}$). La production de la protéine rTrfbL2 (native ou mutée) a été induite avec 0,1 mM d'IPTG lorsque la culture à 37 °C avait atteint une DO_{600} égale à 0,7-0,9. Après induction, les cellules ont été transférées à 16°C sur la nuit. Les cultures ont été centrifugées à 5000g pendant 10 min afin de récupérer les cellules qui ont alors été soit transférées dans un tube de 50 mL pour congélation à -20 °C ou soit traitées directement. Le culot a été resuspendu dans du tampon contenant 20 mM Tris pH 7.5 et 100 mM NaCl à raison de 5 mL de tampon par gramme de cellules obtenu. Les cellules ont ensuite été cassées avec un broyeur cellulaire deux fois à une pression de 1.9 KPa (Constant System Ltd, UK) et centrifugées à froid à 30000 g pendant 30 min. Le surnageant a été filtré comme précédemment et constituait l'extrait soluble.

2.1.4 rPhoNLSeMet

La protéine rPhoNL a également été produite avec des méthionines sélénées (rPhoNLSeMet) en vue de la détermination des phases lors de la résolution de sa structure par diffraction aux rayons X. La méthode décrite par Doublé a été suivie (Doublé, 1997). Les précultures ont été réalisées dans du milieu LB + kanamycine comme pour la protéine sauvage. Après centrifugation, les culots ont été lavés 2 fois avec du milieu M9, mis en culture dans du milieu M9 (2 litres) à 37°C jusqu'à une DO_{600} égale à 0,5 puis placées à 16°C jusqu'à une DO_{600} égale à 0,8-0,9. Les différents acides aminés (Lys, Phe, Thr, Ile, Leu, Val et SeMet à 4 $\text{mg}\cdot\text{mL}^{-1}$) ont alors été ajoutés, et le mélange a été incubé 15 min avant induction avec de l'IPTG à 0,1 mM. Après induction sur la nuit, les cellules ont été centrifugées comme décrit ci-dessus.

2.2 Purification

2.2.1 rPhoSL, rPhoNL, et leurs mutants

Une fois les extraits solubles de rPhoSL et rPhoNL obtenus, la purification des protéines a été réalisée à l'aide d'une colonne d'affinité au nickel 1 ml cOmplete™ His-Tag (Roche). L'équilibration de la colonne a été réalisée avec un tampon contenant 50 mM de Tris pH 8 et 300 mM de NaCl. L'éluion a été réalisée par un gradient linéaire de 0 à 500 mM Imidazole de 20 mL. Après un dessalage sur colonne PD10 (GE Healthcare, Life Sciences) dans un tampon sans NaCl ni imidazole, les étiquettes (histidine + DsbA) ont été coupées par la protéase TEV au 1/50^{ème} sur la nuit à 19°C. Les protéines clivées ont alors été séparées de leurs étiquettes et de la protéase par un nouveau passage sur la colonne d'affinité au nickel. La protéine d'intérêt, ne pouvant plus se fixer à la colonne, est récupérée dans la fraction non

retenue. Les protéines ont ensuite été soumises à une colonne échangeuse d'anions de 1 ml HiTrap DEAE Sepharose FF (GE Healthcare, Life Sciences). L'équilibration a été réalisée avec un tampon 50 mM Tris pH 8 et l'élution avec un gradient de 0 à 1 M NaCl de 30 mL. Les fractions contenant la protéine ont alors été dessalées sur PD10 dans du MES 25 mM pH 6 puis concentrées par centrifugation sur Microsep 3 kDa (PALL) à environ 6 mg.ml⁻¹ et conservées à 4°C.

2.2.2 rTrfbL2-WT, rTrfbL2-Y111Q, rTrfbL2-Y111A, rTrfbL2-R100G

Pour ces mutants, l'extrait soluble a été déposé sur une colonne de GlcNAc agarose de 5 mL (Sigma-Aldrich). La colonne est ensuite lavée avec un tampon 20 mM Tris pH 7,5, 100 mM NaCl (tampon A). Dans le cas de rTrfbL2-WT, un second lavage est réalisé avec 20 mM Tris pH 7,5 et 1.5M NaCl puis la colonne est à nouveau rincée avec du tampon A. La protéine est éluée avec 20 mM Tris pH 7,5, 100 mM NaCl et 200 mM GlcNAc. Elle a ensuite été dialysée dans de l'eau ultrapure puis concentrée sur un Vivaspin20, cutoff 5 kDa (Sartorius).

2.2.3 rTrfbL2-Y111QD77A, rTrfbL2-D77A

Comme les mutants rTrfbL2-Y111QD77A et rTrfbL2-D77A ne se fixent pas sur colonne à GlcNAc, d'autres méthodes de purification ont été testées. L'ensemble des conditions et du matériel utilisé sont décrits dans le *Tableau III-3*.

Technique utilisée	Matériel	Tampon d'équilibration	Tampon d'élution
Chromatographie d'affinité sur résine sucre	5mL D-Galactose Agarose (Thermo Scientific)	20 mM Tris pH 7.5	100 mM Gal
Chromatographie d'affinité sur résine sucre	D-GalNAc Agarose (Sigma-Aldrich)		100 mM GalNAc
Chromatographie d'affinité sur résine sucre	Sepharose 4B (Sigma-Aldrich)		100 mM Gal et 1M NaCl
Chromatographie d'affinité sur métal immobilisé	1 ml cOmplete™ His-Tag (Roche)		20 mM Tris pH 7.5, 300 mM NaCl, 500 mM Imidazole
Chromatographie échangeuse de cations	<ul style="list-style-type: none"> • Hitrap SP-FF 1 mL • Hitrap CM FF 1 mL Hitrap Capto S 1 mL (GE-healthcare Life Sciences)	20 mM Tris pH 7.5 (ou sodium acetate pH 5.5)	20 mM Tris pH 7.5 (ou sodium acetate pH 5.5) +1M NaCl

Chromatographie échangeuse d'anions	<ul style="list-style-type: none"> • Hitrap Q-FF 1 mL • DEAE FF 1 mL (GE-Healthcare Life Sciences) 	20 mM Tris pH 7.5	20 mM Tris pH 7.5 + 1M NaCl
Précipitation au sulfate d'ammonium	Sulfate d'ammonium (90% saturation)	 	
Chromatographie d'interaction hydrophobe	Hitrap Phenyl FF HighSub 1 mL (GE-Healthcare Life Sciences)	20 mM Tris pH 7,5 + 1,5 M sulfate d'ammonium	20 mM Tris pH 7.5
Chromatographie d'affinité sur résine héparine	Heparin HyperD (Biosepra)	20 mM Tris pH 7	20 mM Tris pH 7 + 1M NaCl
		Mise en place	
Choc thermique	Bain marie	50°C 15 min ; 60°C 15 min ; 70°C 15 min → centrifugation : analyse du surnageant / extrait insoluble	

Tableau III-3. Ensemble des conditions testées pour la purification des mutants rTrfbL2-Y111QD77A et rTrfbL2-D77A.

3. Caractérisation biochimique

3.1 Électrophorèse sur gel de polyacrylamide

La technique SDS-PAGE (Electrophorèse en Gel de PolyAcrylamide contenant du dodécylsulfate de sodium) permet de séparer les protéines en fonction de leur taille, sans tenir compte de leurs charges intrinsèques ni de leur structure tridimensionnelle. Il existe deux principales techniques les plus couramment utilisées : la méthode Glycine-SDS-PAGE (ou méthode de Laemmli (Laemmli, 1970)) et la méthode Tricine-SDS-PAGE (Schägger, 2006), dont la différence repose sur l'utilisation de tampon à base de glycine ou de tricine.

3.1.1 Gel SDS-PAGE Tris-glycine

La méthode Glycine-SDS-PAGE permet de séparer des protéines entre 15 et 200 kDa. Les protéines sont préalablement chauffées en présence d'un détergent, le dodécylsulfate de sodium (SDS), qui, en interagissant avec la protéine, va masquer ses charges intrinsèques et former un complexe chargé négativement. Cela va également avoir pour conséquence de dénaturer la protéine et de la solubiliser. Dans le cas où des ponts disulfures sont présents, il peut être nécessaire de compléter cette dénaturation en ajoutant un agent réducteur tel que le 2-mercaptoéthanol ou le DTT (dithiothréitol). La séparation des protéines se fait ensuite sur un gel de polyacrylamide relié à générateur, de façon à faire migrer les protéines chargées

négativement vers l'anode. Ce gel est constitué de deux parties : le gel de concentration à pH 6,8 et le gel de séparation à pH 8.8. Le tampon d'électrophorèse est constitué de Tris-HCl-glycine. La glycine, faiblement chargée à pH 6,8 va migrer très lentement, contrairement aux ions Cl⁻ qui sont beaucoup plus rapides. Les protéines, quant à elles, vont migrer à une vitesse intermédiaire, ce qui va les concentrer entre les fronts de la glycine et des ions Cl⁻ sur une zone restreinte (Smith, 1994 ; Gallagher, 2007). Ainsi, elles vont pouvoir commencer à migrer sur le gel de séparation en même temps, ce qui va permettre d'avoir une résolution plus nette. Les protéines de haut poids moléculaire vont migrer plus lentement que les plus petites protéines car elles sont sujettes à davantage de frottements avec les pores du gel. En fonction des tailles de protéines à séparer, la taille des pores peut être adaptée en modifiant les concentrations d'acrylamide lors de la préparation des gels.

Le protocole utilisé pour la purification des protéines dans le cadre de ce travail de thèse est le suivant : les fractions protéiques ont été préalablement diluées dans un tampon dénaturant constitué de 0.25 M Tris pH 6.8, 50 % glycérol, 10 % SDS, 0.05 % bleu de bromophénol (+ 0.25 M DTT en condition réductrice). Les échantillons sont ensuite chauffés à 100°C pendant 5 min avant d'être déposés sur gel de polyacrylamide 15 %. La migration est réalisée à 200 V pendant 1 heure. Les gels sont ensuite colorés avec de l'« Instant Blue » (Expedeon) pendant 15 min sous agitation.

3.1.2 Gel SDS-PAGE Tris-tricine

Lorsque que l'on travaille avec des protéines de poids moléculaire inférieur à 15 kDa, la méthode de Laemmli n'est plus adaptée. Cela est dû au fait que les petits peptides forment des complexes peptide-SDS de taille et charge très proches des micelles de SDS elles-mêmes et auront donc tendance à être entraînés par ces micelles qui se trouvent au niveau du front de migration. De plus, les peptides vont migrer beaucoup plus vite que la glycine dans le gel de concentration et leur migration sera très aléatoire. On peut alors utiliser la méthode Tricine-SDS-PAGE qui permet de séparer des protéines entre 1 et 100 kDa (Schägger, 2006). La tricine migre plus vite que la glycine pour des pH entre 6.8 et 8.8 car son point isoélectrique est plus faible que celui de la glycine. Elle permet de mieux concentrer puis séparer les protéines de faible poids moléculaire en évitant leur interaction avec les micelles de SDS (Schägger and von Jagow, 1987).

D'autres modifications dans la composition des tampons et des gels est associée. Le détail des préparations utilisées dans cette étude pour des gels Tris-tricine 16.5% est indiqué dans le *Tableau III-4*.

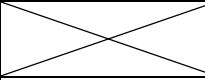
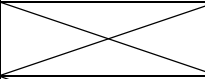
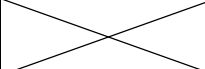
	Gel de séparation	Gel de concentration
Acrylamide (37.5 :1)	16,5 %	3.8 %
Tris pH 8.45	1 M	0.7 M
glycérol	10%	
SDS	0.1%	0.1%
APS	0.08%	0.1%
TEMED	0.03%	0.1%
	Tampon cathode	Tampon anode
Tris	0.1M	0.2 M pH 8.9
Tricine	0.1M	
SDS	0.1 %	

Tableau III-4. Conditions pour la préparation de gels Tris-tricine 16.5 % et des tampons de migration 1X.

Les fractions protéiques ont été suspendues dans le même tampon que pour la méthode Tris-Glycine décrite dans le paragraphe précédent. La migration a été réalisée à 25 mA pendant 2h30 et en utilisant des tampons froids (4°C).

3.2 Chromatographie d'exclusion de taille

La chromatographie d'exclusion de taille, aussi appelée gel filtration, permet de séparer les molécules en fonction de leur taille et de leur forme. Les colonnes de gel filtration contiennent une résine constituée de granules de gel poreux. Les molécules de taille supérieure aux pores sont éluées en premier car elles ne pourront pas diffuser à l'intérieur des billes, elles sont éluées dans ce qu'on appelle le volume mort. Les molécules plus petites vont pénétrer dans les pores à différents degrés selon leur taille. Grâce à une courbe de calibration avec des protéines de poids moléculaire connu, leur volume d'éluion pourra être associé à une masse

moléculaire. Il faut cependant être prudent : cela va donner des valeurs de poids moléculaire précises dans le cas de molécules globulaires, mais les résultats peuvent être source d'erreurs dans le cas de protéines allongées. En effet, ces dernières vont avoir tendance à être éluées plus vite qu'une protéine globulaire de même poids moléculaire et il en résultera une surestimation de sa taille.

La chromatographie d'exclusion de taille peut être utilisée pour différentes applications. Elle permet tout d'abord d'améliorer la purification des molécules en les séparant selon leur poids moléculaire, par fractionnement du volume d'élution. Elle peut aussi être utilisée pour réaliser des études analytiques d'un échantillon déjà purifié. Associée à un détecteur de fluorescence ou à un appareil de diffusion de la lumière à plusieurs angles (MALS : Multiangle Light Scattering), elle va permettre de vérifier la qualité, la pureté, ou les propriétés d'un échantillon. Par exemple, cette technique est couramment utilisée pour déterminer l'état d'oligomérisation des protéines. Enfin, la gel filtration est un outil pratique pour le dessalage et le changement de tampon des échantillons. En effet, les petites molécules telles que le sel ou autre composants des tampons vont être éluées après les molécules de plus grande taille, telles que les protéines. Les colonnes PD10 évoquées *chapitre 2.2.1* sont un exemple de colonnes d'exclusion de taille utilisées pour le dessalage.

Dans cette étude, les colonnes d'exclusion de taille Enrich 70 (BioRad) et S200 (GE Healthcare) qui permettent de séparer les protéines de tailles comprises entre 0.5-70 kDa et 10-600 kDa respectivement ont été utilisées. La colonne Enrich 70 a été calibrée avec la BSA (66.5 kDa), l'ovalbumine (43 kDa), et la ribonucléase A (13.7 kDa).

3.3 SEC-MALS

La méthode SEC-MALS est la combinaison d'une chromatographie d'exclusion de taille (SEC) avec un appareil de détection multi angulaire de la diffusion de la lumière (MALS) associé à un réfractomètre différentiel (dRI). Elle permet la détermination précise de la masse molaire absolue et du rayon de giration (R_g) des biomolécules au cours de leur séparation sur la colonne de chromatographie. Cette méthode se base sur le fait que la diffusion de la lumière par les molécules est proportionnelle au produit de sa concentration et de sa masse moléculaire. Le réfractomètre différentiel permettant de mesurer directement la concentration des biomolécules, la mesure de la diffusion de la lumière à différents angles permet de déduire la masse moléculaire. Cette méthode est beaucoup plus précise que la chromatographie d'exclusion de taille où les résultats peuvent être biaisés par la forme des

biomolécules ou leur interaction non-spécifique avec la résine de la colonne utilisée. Les mesures SEC-MALS ont été réalisées au CEA de Grenoble (Laboratoire Chimie et Biologie des Métaux).

3.4 Diffusion dynamique de la lumière (DLS)

La diffusion dynamique de la lumière est une technique rapide et non-invasive permettant d'estimer le poids moléculaire des molécules en solution. Appliquée aux protéines, elle permet aussi de déterminer leur état d'oligomérisation, de s'assurer de la monodispersité et de l'homogénéité d'un échantillon ou encore de vérifier l'absence d'aggrégation.

Cette technique repose sur la mesure des fluctuations de l'intensité de la lumière diffusée en fonction du temps. Un faisceau laser est dirigé dans la solution contenant les molécules à analyser. Quand ce laser rencontre une particule, l'énergie lumineuse va être dispersée dans toutes les directions et la lumière ainsi diffusée sera analysée sur un détecteur à photons relié à un corrélateur. Le fait que les molécules soient soumises au mouvement brownien, va entraîner des fluctuations de la lumière diffusée. Plus les molécules sont petites, plus leur mouvement brownien sera rapide, entraînant des fluctuations d'intensité rapide. Ainsi, les informations de fluctuations d'intensité vont permettre de déterminer le rayon hydrodynamique de l'objet diffusé. L'appareil utilisé lors de cette étude est le zétasizer nano series (Malvern Panalytical).

3.5 Test de stabilité thermique

Le test de stabilité thermique (TSA pour « Thermal Shift Assay), est une technique qui permet d'étudier la stabilité des protéines vis-à-vis de la température et de rechercher les facteurs qui peuvent influencer cette stabilité (tampon, présence de sels, présence d'un ligand...). Le principe repose sur l'utilisation d'une sonde hydrophobe fluorescente et dans notre cas, il s'agit du sypro-orange (Sigma Aldrich). Cette sonde ne fluoresce que dans un environnement hydrophobe. Plus les protéines vont se déplier sous l'effet de la température, plus elles vont exposer leurs régions hydrophobes et donc plus de molécules de la sonde fluorescente vont avoir tendance à se fixer. En suivant le taux de fluorescence en fonction de la température (augmentation de 1 degré par min de 25 à 100°C), on peut en déduire la température de fusion T_m . Elle est définie comme la température à laquelle on observe un équilibre entre les populations de protéines repliées et dépliées. Le test s'effectue dans des plaques 96 puits avec un appareil de QRT-PCR (Mini-Opticon, Biorad). Différent tampons à

différents pH, en absence ou présence de sel, ont été testés avec les protéines rPhoSL, rPhoNL, et rTrfbL2.

4. Marquage des lectines

4.1 Alexa 488, Alexa 680 et biotine

Plusieurs tests décrits par la suite ont nécessité le marquage des lectines avec des fluorophores. Les lectines ont été marquées à l'Alexa Fluor 488 NHS Ester (Ester de N-hydroxysuccinimide, Fisher Scientific), Alexa Fluor 680 NHS Ester (Ester de N-hydroxysuccinimide, Fisher Scientific), et Biotine (acide biotinamidohexanoyl-6-aminohexanoïque, Ester de N-hydroxysuccinimide; Sigma Aldrich) selon les recommandations du fournisseur. Le tampon des protéines a été échangé pour du bicarbonate de sodium à pH 8,3 sur colonne PD10 avant marquage. Les solutions d'Alexa ou de biotine sont incubées avec la protéine avec un rapport de 1 mg de marqueur pour 10 mg de protéines pendant 1 h sous faible agitation. Puis la protéine marquée a été séparée du conjugué non lié sur colonne PD10 équilibrée avec le tampon de conservation de la lectine.

4.2 Peroxydase

Dans le but de réaliser des études de lectine-histochimie sur tissus cancéreux, les lectines ont aussi été marquées à la peroxydase. Pour cela, le kit EZ-Link Plus Activated Peroxidase (Thermo_Fisher) a été utilisé. Les lectines ont été reprises dans le tampon Carbonate-Bicarbonate pH 9.4 fournit avec le kit. Le marquage a été réalisé en ajoutant 1 mg de EZ-Link Plus Activated Peroxidase pour 1 mg de lectine. Le mélange a été incubé 1 heure avant d'ajouter 10 µL de cyanoborohydrure de sodium. Après 15 min d'incubation, 20 µL de Quenching Buffer ont été ajoutés et le mélange incubé à nouveau 15 min. Toutes les incubations ont été réalisées à température ambiante. Le tampon a été ensuite changé pour du MES pH 6 à 25 mM sur colonne PD10, et les lectines ont été conservées à 4°C ou à -20°C. rPhoNLm a également été marquée indirectement avec de la streptavidine-peroxydase (sigma) après avoir été biotinylée.

5. Étude de l'interaction lectines-sucres

5.1 Tests d'hémagglutination de globules rouges de lapin

Le test d'hémagglutination est une méthode rapide de première caractérisation des lectines (Golstein et al., 1980). Les lectines sont généralement des protéines multivalentes, c'est-à-dire qu'elles possèdent plusieurs sites de reconnaissance par molécule. Cette particularité leur

permet d'héماغglutiner les globules rouges qui possèdent les ligands reconnus par la lectine à leur surface. En effet, en se fixant à ces ligands sur au moins deux globules rouges à la fois, elles vont contribuer à la formation d'un réseau de cellules. En absence d'héماغglutination, au contraire, les globules rouges ne forment pas de réseau et sédimentent au fond des puits. Le test d'héماغglutination peut avoir différents buts : vérifier l'activité des lectines, autrement dit leur capacité à reconnaître leur ligand ; vérifier leur multivalence, qui peut être lié à leur état d'oligomérisation ; ou encore étudier les ligands qu'elles pourraient reconnaître en réalisant une inhibition de l'héماغglutination avec différents sucres.

Les tests d'héماغglutination ont été réalisés par séries de dilutions au ½ de l'extrait protéique dans 150 mM de NaCl pour un volume final de 25 µl. Le test a été réalisé sur des microplaques de titration 96 puits en U (NunC) en ajoutant 25 µl de globules rouges de lapin 3% (Biomérieux) dans 150 mM de NaCl. Dans certaines conditions, les globules rouges ont été préalablement traités à la trypsine pour permettre la fixation de la lectine sur des ligands difficiles d'accès. Pour cela, après dilution à 3% des globules rouges, la trypsine a été ajoutée à 0,5 mg.mL⁻¹. Après un repos à 37°C pendant 30 min, le mélange a été centrifugé à 2000 g pendant 5 min puis le culot a été lavé 6 fois avec 150 mM de NaCl. Les plaques sont incubées 20 min à 37°C avant d'être analysées. La concentration minimale en lectines permettant une agglutination évidente est ainsi déterminée.

5.2 Analyse de la spécificité sur puces à sucres (glycan array)

Les puces à sucres sont des outils permettant d'analyser les sucres reconnus par un composé donné, comme par exemple des lectines. Dans le but d'analyser les lectines rPhoSL, rPhoNL et rTrfbL2 sur puce à sucre, ces protéines ont été préalablement marquées à l'Alexa Fluor 488-NHS. Elles ont ensuite été analysées sur une puce contenant 600 glycanes de mammifères en suivant la procédure standard du Consortium for Functional glycomics (USA). rPhoSL et rPhoNL ont été analysées à 50 µg.mL⁻¹ et 5 µg.mL⁻¹ sur la puce mammifères version 5.2. rTrfbL2 a été analysée à 0.2 µg.mL⁻¹ sur la puce mammifères version 5.1.

5.3 Tests d'interaction ELLA (Enzyme-Linked Lectin Assay)

Les tests ELLA permettent d'étudier la reconnaissance entre des glycoprotéines et des lectines. Le principe du test qui a été réalisé est décrit *Figure III-4*. Ce test a été réalisé en plaque 96 puits. Les glycoprotéines d'intérêt (ovalbumine, conalbumine et thyroglobuline (GE healthcare), fétuine et IgG1-κ humaines (Sigma Aldrich)) ont été déposées au fond des

puits à raison de 1000 ng par puits et incubées 2 heures à température ambiante. Les glycoprotéines non fixées ont été éliminées par lavage avec du PBS-Ca (137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 1.76 mM KH₂PO₄, 0.1 mM CaCl₂ pH 7.4). Puis les puits ont été recouverts avec 2% de BSA pendant 1 h à 37°C et lavés trois fois avec du PBS-T (PBS-Ca + 0.05% Tween 20). Les protéines, préalablement biotinylées, ont alors été ajoutées à raison de 30 ng de lectine par puits et incubées 2 h à température ambiante. Les plaques ont été lavées avec du PBS-T avant de procéder à l'étape de révélation en utilisant 10 ng par puits de streptavidine-peroxidase (Sigma-Aldrich). Après lavage des plaques 5 fois avec du PBS-T, la couleur a été développée avec du dichlorhydrate d'o-phénylènediamine (SigmaFast OPD tablet, Sigma Aldrich) pendant quelques minutes dans le noir avant de stopper la réaction avec de l'HCl 3 M. Le signal a ensuite été lu par un lecteur de plaques (TECAN, Spark 10M), à 492 nm. Toutes les mesures ont été réalisées en duplicats.

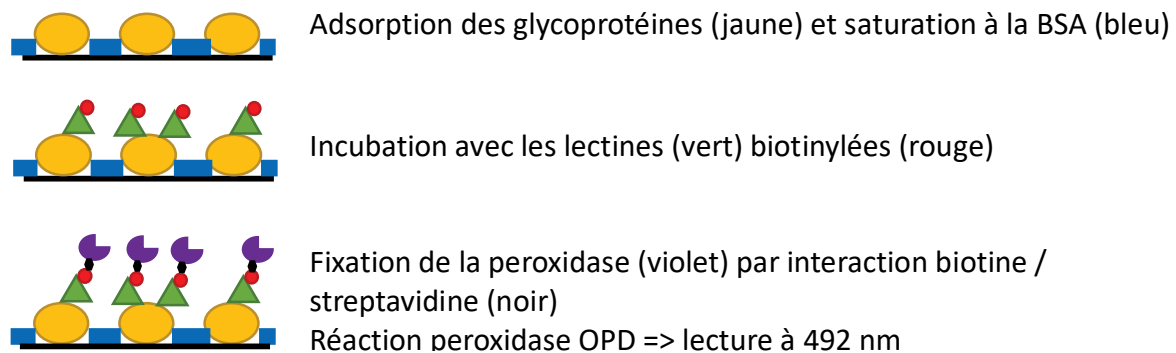


Figure III-4. Principe du test ELLA.

5.4 Titration calorimétrique isotherme (ITC)

La titration calorimétrique isotherme (ITC) est une technique précise permettant d'étudier les interactions entre biomolécules en solution (Dam et al., 2016). Le principe repose sur le fait que ces interactions entraînent des dégagements ou des absorptions de chaleur, selon que la réaction est endothermique ou exothermique. Étudier les interactions consiste donc à mesurer les changements de température. Pour cela, l'appareil est équipé de deux cellules : une qui contient la première biomolécule partenaire (cellule échantillon) et l'autre qui contient uniquement de l'eau (cellule de référence). Le deuxième partenaire de l'interaction est placé au départ dans une seringue et est ensuite injecté dans la cellule par petits volumes, en plusieurs fois (*Figure III-5*). L'interaction entre les partenaires va provoquer des changements de température de la cellule échantillon par rapport à la cellule de référence. Pour maintenir la température constante entre les deux cellules, l'appareil va fournir un

courant électrique. La puissance de ce signal, exprimée en $\mu\text{cal}/\text{sec}$, est enregistrée pour chaque injection et est directement proportionnelle à la chaleur émise ou absorbée durant l'interaction. Au fil des injections, de plus en plus de molécules partenaires entrent en interaction, jusqu'à obtenir une saturation des sites de fixation à la fin de la titration.

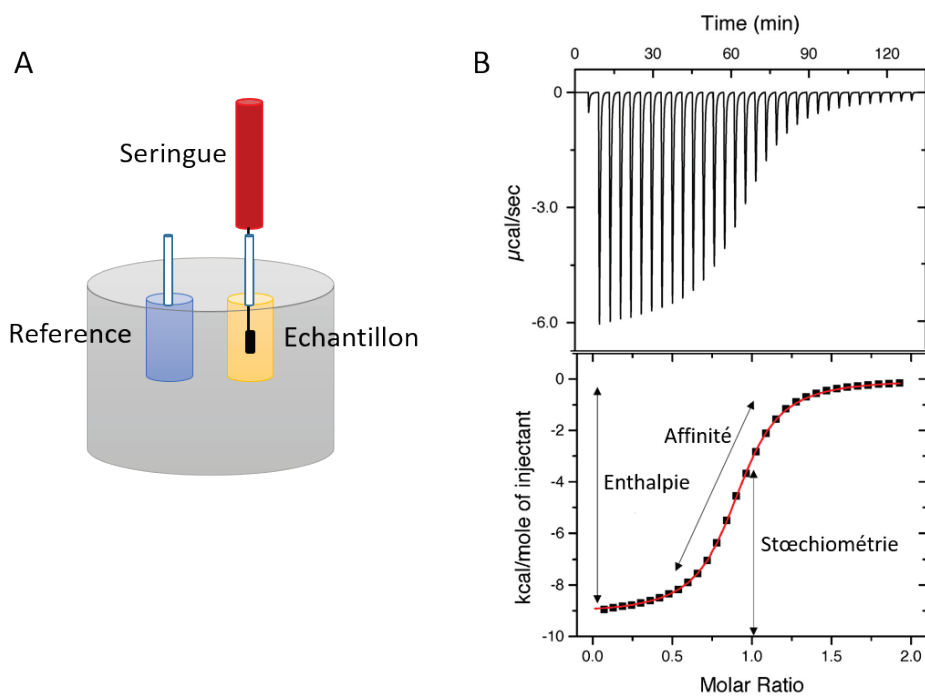


Figure III-5. Schéma du fonctionnement d'un appareil d'ITC (A) et d'un thermogramme classique (B).

A. La cellule de référence contient de l'eau. La cellule échantillon contient un des partenaires d'interaction. La seringue contient le second partenaire qui est injecté de façon répétée et régulière dans la cellule échantillon. B. Le thermogramme donne des informations sur l'enthalpie, l'affinité et la stœchiométrie de la réaction.

Les mesures d'ITC permettent d'obtenir plusieurs informations : la première est la constante d'association à l'équilibre K_A , exprimée en M^{-1} , et qui traduit la force de l'interaction entre les molécules. Généralement on utilise son inverse, la constante de dissociation à l'équilibre K_D (en M), pour exprimer l'affinité. L'ITC permet aussi de mesurer la stœchiométrie n , et l'enthalpie de fixation ΔH . Les autres paramètres thermodynamiques (variation d'énergie libre ΔG et l'entropie ΔS) sont calculés à partir de l'équation suivante :

$$\Delta G = \Delta H - T\Delta S = -RT \ln K_A \text{ (équation 1)}$$

Avec T étant la température absolue en Kelvins et R la constante des gaz parfaits ($R = 8.314 \text{ J mol}^{-1} \text{ K}^{-1}$).

Les mesures d'ITC ont été réalisées sur un microcalorimètre ITC200 (Microcal Inc., Malvern Panalytical). La protéine a été introduite dans la cellule de mesure d'un volume de

200 μl à 25°C. Sa concentration a été vérifiée au préalable avec un spectrophotomètre NanoDrop 1000 (Ozyme) en utilisant le coefficient d'absorption molaire (théorique) de la protéine d'intérêt. La titration a été réalisée avec 20 injections du ligand à raison de 2 μl toutes les 3 minutes. Les données ont été traitées avec le logiciel Origin 7.

5.5 Résonance plasmonique de surface (SPR)

La résonance plasmonique de surface est aussi une méthode de mesure en temps réel des interactions entre biomolécules. Mais contrairement à l'ITC, les partenaires ne sont pas tous les deux libres en solution, puisque l'un d'entre eux est fixé sous forme de monocouche sur une puce. Le principe est d'envoyer une lumière polarisée monochromatique sur la puce, recouverte d'une couche métallique d'or. Une partie de la lumière est réfléchi sur la puce et une onde évanescente (composante électromagnétique de la lumière) va se propager perpendiculairement à travers la puce. Les électrons libres de la couche d'or vont, eux, entrer en résonance avec les photons du faisceau incident, entraînant une chute d'intensité du faisceau réfléchi à son angle de résonance (**Figure III-6**). Cet angle de résonance va varier selon l'indice de réfraction du milieu et donc lors des phases d'association et de dissociation entre les deux partenaires. L'enregistrement au cours du temps de la variation de l'angle de résonance va alors permettre de suivre la fixation des molécules, sous forme d'un sensorgramme des unités de résonances (RU) en fonction du temps. 1000 RU correspond à un déplacement de l'angle de 0.1° et équivaut à la fixation de 1 ng d'analyte par mm^2 .

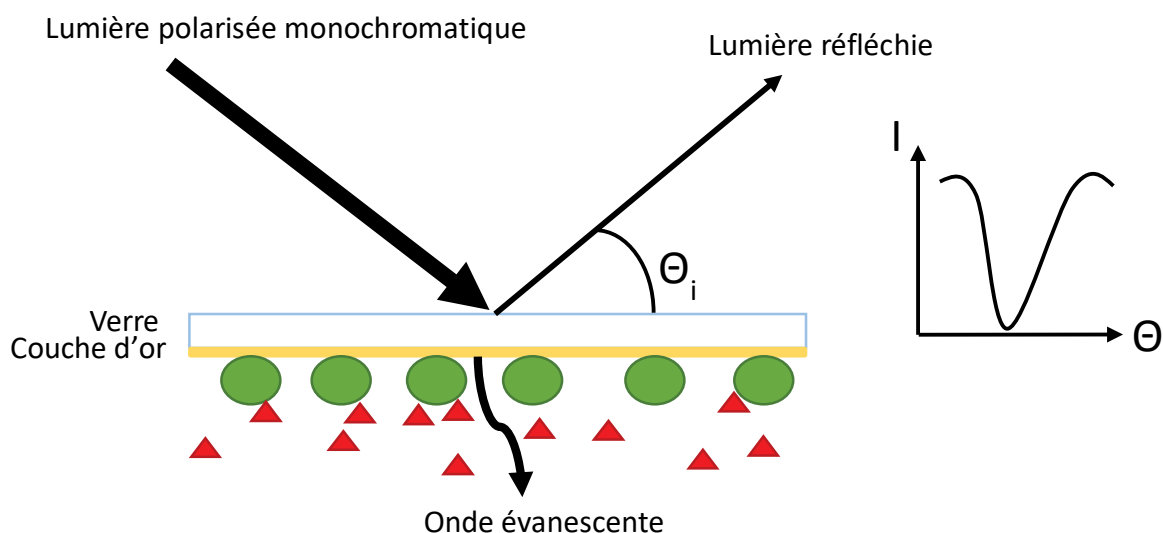


Figure III-6. Principe de la résonance plasmonique de surface (SPR).

En vert, le partenaire fixé à la puce, en rouge le second partenaire en solution. L'appareil mesure les variations de l'angle de résonance au cours du temps, qui varie lors de la fixation entre les deux partenaires.

Les résultats sont ensuite traités en comparant les données avec des courbes théoriques représentant des modèles de liaison moléculaire connus. Cela permet de déterminer les constantes d'affinité (K_D en M) et les constantes cinétiques : constantes de vitesse d'association (k_a en $M^{-1}s^{-1}$) et de dissociation (k_d en s^{-1}).

Dans cette étude, des puces CM5 ont été utilisées. Elles sont recouvertes de dextran carboxyméthylé. L'utilisation d'un mélange d'1-éthyl-3-(3-diméthylaminopropyl) carbodiimide (EDC) et de N-hydroxysuccinimide (NHS) permet d'activer les groupements carboxyles et d'immobiliser de la streptavidine par liaison amide. On sature ensuite les groupements carboxyles restés libres avec de l'éthanolamine avant de fixer les ligands biotinylés.

Cette technique a été utilisée pour mesurer l'interaction entre rPhoSL et la thyroglobuline (qui possède des ligands de rPhoSL a sa surface). Pour cela, la thyroglobuline a été préalablement biotinylée, puis fixée sur le premier canal d'une puce CM5 préalablement fonctionnalisée avec de la streptavidine. rPhoSL a ensuite été ajoutée à différentes concentrations (62.5, 100, 125, 200, 250, 300, 350 et 400 μ M). Le deuxième canal de la puce sert de référence et a juste été fonctionnalisé avec de la streptavidine, sans ajout de thyroglobuline. L'analyse a été réalisée avec un Biacore X100 (GE Healthcare, Life Sciences) dans un tampon phosphate PBS pH 7,4 (Na_2HPO_4 10 mM, KH_2PO_4 1,76 mM, KCl 2,7 mM et NaCl 137 mM).

5.6 Bio-Layer Interferometry (BLI)

La technique de BLI est une technique récente de mesure des interactions entre macromolécules, décrite pour la première fois en 2005 par ForteBio. Le principe est d'envoyer une lumière blanche sur un capteur constitué de deux couches : la couche biocompatible où est fixé l'un des partenaires et une couche interne qui sert de référence. Chaque couche va réfléchir la lumière blanche et les deux réflexions vont interférer entre elles de façon constructive ou destructive selon la longueur d'onde. Ces interférences sont analysées par un spectromètre qui va donner un profil des intensités de lumières réfléchies en fonction de la longueur d'onde. Lors de l'interaction, plus on ajoute de ligand, plus l'épaisseur de la couche biocompatible augmente et donc plus la distance entre les deux couches augmente. Cela se traduit par des décalages du profil d'interférence par rapport à la condition sans ligand (*Figure III-7*). Le profil spectral va donc changer en fonction du nombre de molécules fixées sur la surface du capteur, ce qui permet de suivre en temps réel l'interaction

entre les biomolécules. La mesure de l'ampleur du décalage des profils au cours du temps donne les courbes classiques d'association et de dissociation, qu'on appelle sensorgramme (**Figure III-7.C**). La technique BLI a des similitudes avec la SPR puisqu'on a là aussi un des partenaires qui est fixé sur une surface et l'autre partenaire en solution. Par contre, contrairement à la SPR, on ne travaille pas en flux continu.

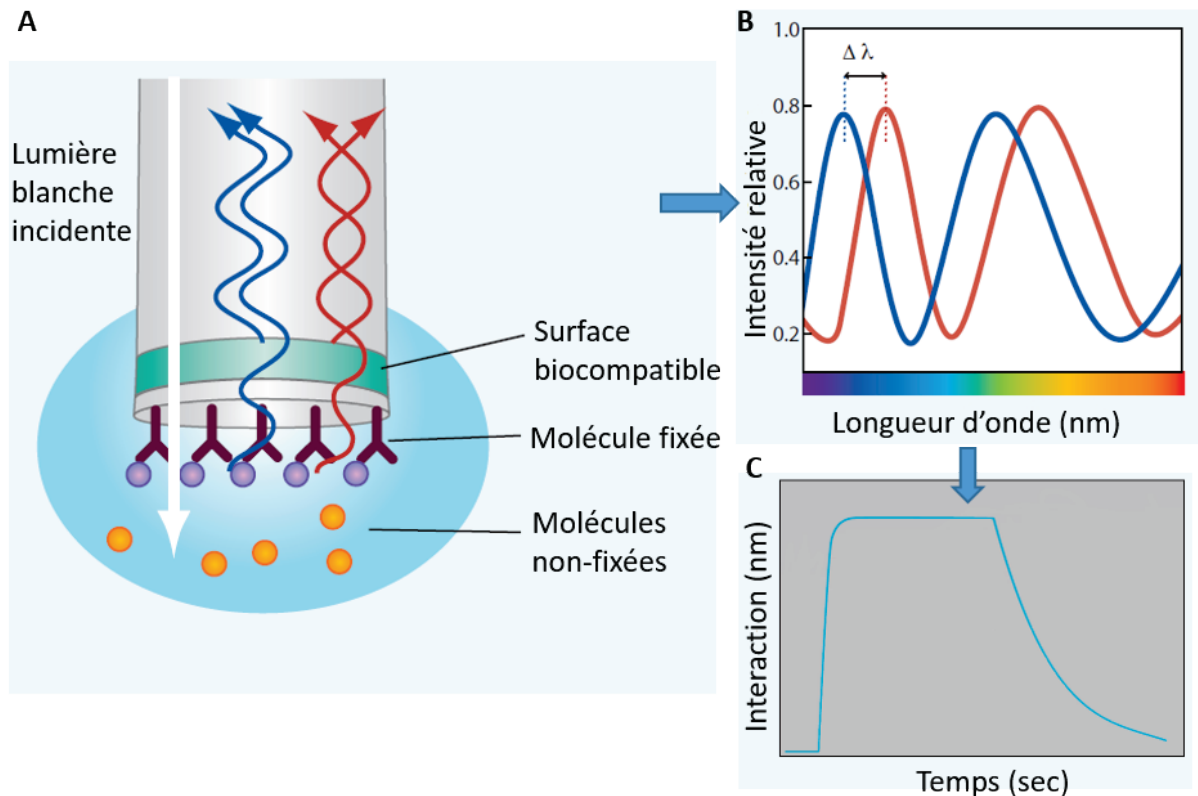


Figure III-7. Principe du BLI.

La lumière blanche est réfléchiée à la fois par la surface biocompatible et par la couche interne (A). La fixation des molécules lors de l'interaction entraîne un décalage du profil d'interférence (B). Sensorgramme classique d'une interaction présentant les phases d'association et de dissociation (C). Figure traduite de (Tobias and Kumaraswamy, 2014).

Cette technique a été utilisée pour mesurer l'interaction entre rPhoSL et la thyroglobuline. Les mesures ont été réalisées sur un appareil OctetRed96 (Forte Bio – PALL) à 25°C dans un tampon PBS pH 7,4. La thyroglobuline biotinyllée a soit été fixée sur des biocapteurs fonctionnalisés avec de la streptavidine (Forte Bio – PALL), soit fixée directement sur des biocapteurs sans streptavidine, par liaison amide. Un biocapteur sans thyroglobuline a servi de référence. Ces capteurs ont ensuite été mis en contact avec différentes concentrations de rPhoSL (62.5, 90, 125, 250 et 300 μM). Dans le but de limiter les interactions non spécifiques entre rPhoSL et le capteur de référence fonctionnalisé avec de la streptavidine, les protéines biotinyllées LecA ou LecB ou du PAA-GlcNAc ont été ajoutées sur le capteur.

6. Étude structurale

6.1 N-glycanes synthétisés pour l'étude

Plusieurs N-glycanes fucosylés en α 1,6 ont été synthétisés dans le cadre de notre collaboration avec l'équipe de Carlo Unverzagt (Université de Bayreuth, Allemagne). Ces synthèses avaient pour objectif d'enrichir les données d'interaction entre différents ligands et rPhoSL et rPhoNL. L'ensemble des N-glycanes en question sont présentés schématiquement **Figure III-8** (voir les détails des structures des sucres en Annexe **Tableau VII-2**). Les codes des différents N-glycanes proviennent du site GlyTouCan (<https://glytoucan.org/>, (Aoki-Kinoshita et al., 2016)) qui permet d'associer à chaque structure un identifiant unique.

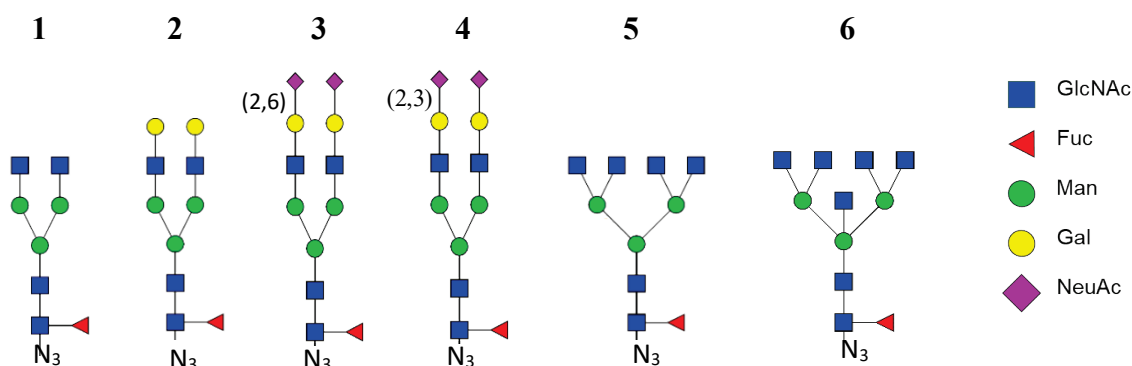


Figure III-8. Ensemble des N-glycanes fucosylés en α (1,6) utilisés dans l'étude.

N-glycane 1 (G80858MF), N-glycane 2 (G78059CC), N-glycane 3 (G56749GV), N-glycane 4 (G17689DH), N-glycane 5 (G00176HZ), N-glycane 6 (G40358EZ). Entre parenthèses, sont indiqués les codes GlyTouCan.

6.2 Principes de la cristallographie aux rayons X

La cristallographie aux rayons X est une technique permettant de déterminer la structure atomique de molécules en se servant du principe physique de diffraction des rayons X par la matière. Elle nécessite une première étape de cristallisation des molécules d'intérêt.

6.2.1 La cristallogénèse

La formation d'un cristal est un processus complexe, dont les conditions optimales sont impossibles à prédire car les phénomènes de solvatation et de précipitation ne sont pas assez bien compris. Il s'agit d'un cas particulier d'agrégation ordonnée et répétitive. Pour trouver les bonnes conditions de cristallisation, il faut donc tester un large panel de conditions différentes en se basant sur des mélanges ayant déjà conduit à la formation de cristaux pour d'autres molécules. Plusieurs méthodes de cristallogénèse sont possibles, la technique de

diffusion de vapeur en goutte suspendue ou en goutte assise étant la plus couramment utilisée en cristallisation manuelle (**Figure III-9**). Le principe de la goutte suspendue est de former, sur une lamelle, une goutte contenant généralement 50 : 50 d'agent cristallisant et de solution protéique. Cette goutte est ensuite retournée sur un réservoir contenant de l'agent de cristallisation et scellée avec de la graisse. La concentration en agent cristallisant est deux fois moins élevée dans la goutte que dans le réservoir, ce qui va entraîner des échanges de vapeurs.

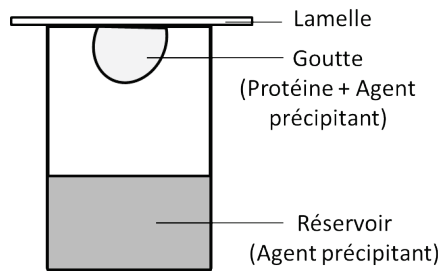


Figure III-9. Cristallogénèse : technique de diffusion de vapeur en goutte suspendue.

Le diagramme de phase permet d'expliquer la physico-chimie du phénomène de cristallisation qui est fonction des concentrations en protéine et en agent cristallisant (**Figure III-10**). Les échanges gazeux entraînent une augmentation des concentrations en macromolécule et en agent précipitant de façon à traverser la zone métastable et atteindre la zone de nucléation où se forment des noyaux stables. Il y aura ensuite cristallisation si la concentration en protéine soluble de la goutte converge vers la courbe de solubilité.

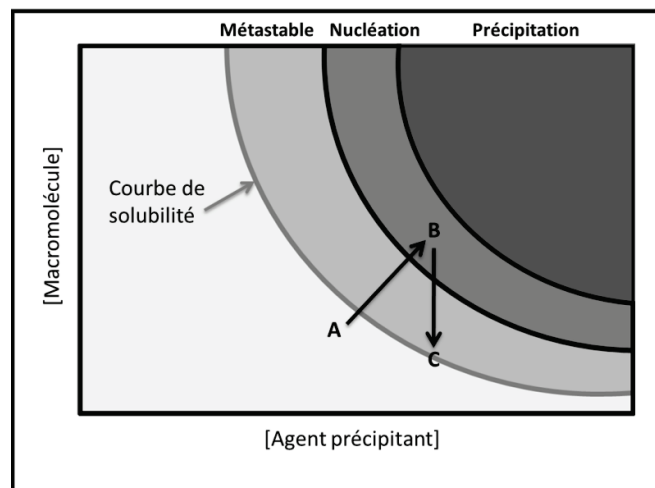


Figure III-10. Diagramme de phase et représentation des conditions nécessaires à la cristallisation par la technique de diffusion de vapeur.

A : molécule soluble ; B : zone de formation de noyaux de cristallisation ; C : limite de croissance cristalline.

6.2.2 Phénomène de diffraction des rayons X

La technique de cristallographie aux rayons X repose sur le phénomène de diffusion de la lumière qui est également à la base de l'imagerie optique. En effet, en microscopie optique, la lumière est diffusée par chaque point de l'objet éclairé, les rayons étant ensuite convergés par des lentilles afin de reconstituer l'image. Pour observer un objet, on a besoin d'une lumière de longueur d'onde adaptée à la taille de l'objet. De ce fait, pour observer des détails à résolution atomique, on utilise des rayons X, qui ont une longueur d'onde de l'ordre de l'Ångström. Cela implique néanmoins que l'on ne peut pas utiliser de lentille pour refocaliser les rayons X, car il n'en existe aucune capable de le faire. La diffraction aux rayons X est donc une technique d'imagerie sans lentille où les ondes diffusées, non refocalisées, appartiennent à l'espace réciproque (en opposition à l'espace réel). Un traitement informatique et mathématique basé sur la transformée de Fourier permet ensuite de passer de l'espace réciproque à l'espace réel de façon à reconstruire une image de la molécule étudiée.

Les rayons X interagissent très peu avec la matière, il n'est donc pas possible d'utiliser cette méthode sur une seule molécule. C'est la raison pour laquelle on travaille avec des cristaux, qui sont des arrangements périodiques ordonnés en trois dimensions de molécules. Cela permet d'augmenter le nombre de diffuseurs, d'augmenter la résistance au rayonnement x et donc d'obtenir un meilleur signal. De plus, les molécules du cristal sont dans la même orientation, ce qui favorise les interférences constructives que l'on va pouvoir observer sur le cliché de diffraction (*Figure III-11*).

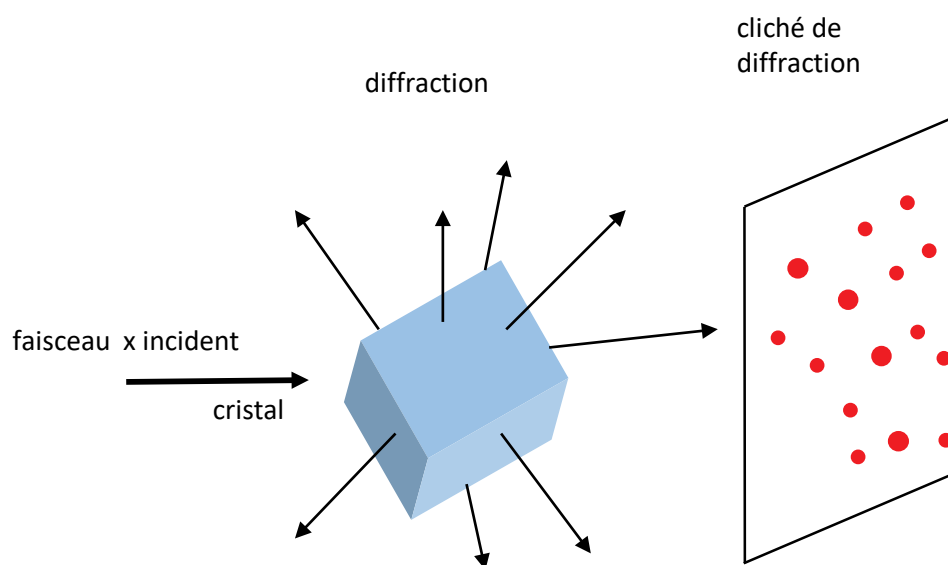


Figure III-11. Diffraction des rayons X par un cristal à l'origine du cliché de diffraction.

Le plus petit groupe d'atomes permettant de reconstituer le cristal par translation dans les 3 dimensions est appelé maille élémentaire. La maille peut contenir une ou plusieurs copies de la molécule d'intérêt. Elle est définie par trois vecteurs a , b et c et trois angles α , β , et γ . Chaque cristal peut être décrit par son système cristallin et son type de réseau, les différentes combinaisons possibles définissent les quatorze réseaux de Bravais (*Tableau III-5*). L'application d'éléments de symétrie à ces réseaux forme 230 groupes d'espaces. Seuls 65 groupes d'espaces sont possibles pour les protéines car elles ne sont pas centrosymétriques : pas de point d'inversion ou de plan miroir. L'unité asymétrique de la maille est le plus petit volume qui permet de reconstruire la structure cristalline en appliquant toutes les opérations de symétrie du groupe d'espace (translations et rotations selon différents axes).

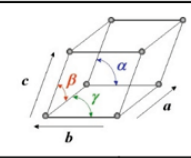
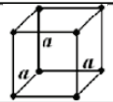
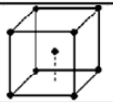
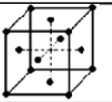
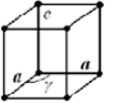
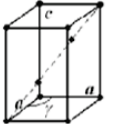
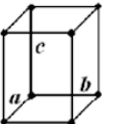
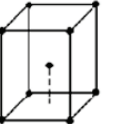
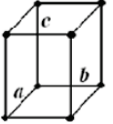

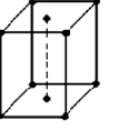
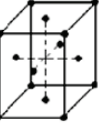
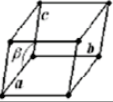
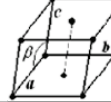
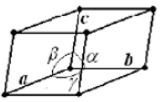
Maille élémentaire Définition des vecteurs a , b , c et des angles α , β , γ					
Système cristallin	Réseaux de Bravais				Exemples de groupes d'espaces
	P	I	C	F	
Cubique $a = b = c$ $\alpha = \beta = \gamma = 90^\circ$					P23, P4 ₁ 32, I432 ...
Trigonal/Hexagonal $a = b \neq c$ $\alpha = \beta = 90^\circ \gamma = 120^\circ$					P6, P6 ₂ , P6 ₃ 22 ...
Trigonal Rhomboédrique $a = b = c$ $\alpha = \beta = \gamma \neq 90^\circ$					P3, P3 ₁ 2, R32 ...
Tétragonal ou quadratique $a = b \neq c$ $\alpha = \beta = \gamma = 90^\circ$					P4, P4 ₂ 2, I4 ₁ 22 ...
Orthorhombique $a \neq b \neq c$ $\alpha = \beta = \gamma = 90^\circ$					P222, P2 ₁ 2 ₁ 2 ₁ , P2 ₁ 2 ₁ 2 ...
Monoclinique $a \neq b \neq c$ $\alpha = \beta = 90^\circ \neq \gamma$					P2, P2 ₁ , C2
Triclinique $a \neq b \neq c$ $\alpha \neq \beta \neq \gamma$					P1

Tableau III-5. Description de la maille élémentaire, des sept systèmes cristallins, des 14 réseaux de Bravais et quelques exemples de groupes d'espaces possibles pour chaque système. P : Primitif, I : corps centré, C : bases centrées, F : Faces centrées.

Les rayons X sont diffusés par le nuage électronique des atomes et l'amplitude de la diffusion est fonction du numéro atomique de cet atome. En cristallographie aux rayons X, l'objectif est de reconstruire la carte de densité électronique des molécules à partir du cliché de diffraction sachant que chaque tache du cliché contient de l'information sur la densité électronique en tout point de la maille cristalline. La distribution des taches de diffraction permet de retrouver le système cristallin (vecteurs et angles). Chaque tache du cliché de diffraction est caractérisée par son facteur de structure F_{hkl} qui renferme deux informations essentielles : l'amplitude de l'onde diffractée (que l'on peut calculer à partir des intensités I_{hkl} de chaque tache) qui permettra de retrouver la nature des atomes et la phase de l'onde (ou déphasage par rapport au rayon incident) qui permettra de retrouver la position des atomes. L'intensité des taches est facilement mesurable avec un détecteur, mais l'information de la phase ne peut pas être enregistrée et est donc perdue. Dans le cas de petites molécules, des méthodes directes de calcul des phases sont possibles, mais dans le cas de macromolécules, la complexité devient limitante.

6.2.3 Le problème des phases

Pour remédier au problème de la perte d'informations des phases, plusieurs méthodes peuvent être utilisées. Lorsque l'on a une structure homologue disponible (minimum 20% d'homologie de séquence), on peut utiliser la méthode de remplacement moléculaire. Traditionnellement, elle est basée sur la méthode de Patterson et consiste à orienter et positionner le modèle dans la maille de la structure inconnue de manière à ce que le profil de diffraction calculé du modèle soit le plus proche possible du profil de diffraction observé. L'orientation du modèle se fait en deux étapes avec d'abord l'étape de rotation où l'on détermine l'orientation relative du modèle par rapport à la molécule de structure inconnue, puis l'étape de translation qui permet de superposer les deux molécules préalablement orientées de la même manière. Le principe de base du remplacement moléculaire repose sur la fonction de Patterson. Elle permet de générer une carte de Patterson qui est une carte des vecteurs interatomiques calculée à partir du carré de l'amplitude du facteur de structure et qui ne requiert pas d'informations de phase. Elle comporte des pics pour chaque vecteur positionné entre deux atomes et dont les intensités sont le résultat du produit de la densité électronique de chaque atome. En comparant les pics de Patterson du modèle avec ceux du profil de diffraction, on peut alors calculer les solutions des fonctions de rotation et de translation permettant de positionner le modèle dans la maille de la structure inconnue. On

utilise ensuite les informations de phases du modèle correctement positionné pour déterminer celles de la structure inconnue.

En absence de structure homologue, on peut recourir aux méthodes de substitution isomorphe simple SIR (Single Isomorphous Replacement) ou multiple MIR. Elles consistent à comparer les données de diffraction I_{hkl} d'un cristal natif aux données de cristaux dans lesquels ont été ajoutés des atomes lourds (un seul cristal pour la méthode SIR et plusieurs pour la méthode MIR). L'isomorphisme entre les cristaux doit être maintenu, c'est-à-dire que les atomes lourds ne doivent pas modifier les paramètres de maille. Ces atomes ont un nombre élevé d'électrons et donc la forte intensité de leurs pics de diffraction diffère significativement de celles de la molécule d'intérêt. La sous-structure des atomes lourds, bien moins complexe que la structure de la macromolécule, peut alors être déterminée avec les méthodes de phasage qui utilisent la carte de Patterson. On calcule alors la contribution de l'atome lourd sur chaque facteur de structure et on en déduit les phases de la macromolécule d'intérêt.

Il est également possible d'utiliser des méthodes de diffusion anormale simple SAD (Single wavelength Anomalous Scattering) ou multiple MAD (Multi-wavelength Anomalous Scattering). La méthode SAD a pour avantage de s'affranchir des problèmes d'isomorphisme en utilisant un seul cristal. Ce cristal est modifié par ajout d'un diffuseur anormal par trempage des cristaux dans des atomes lourds ou synthèse de la protéine avec des acides aminés modifiés tels que les méthionines sélénées (Se-Met) par exemple. Les diffuseurs anormaux ont un facteur de diffusion qui varie fortement dans la région de leur seuil d'absorption. On va donc enregistrer les données de diffraction dans la région du seuil d'absorption du diffuseur anormal et ensuite retrouver les phases par méthode directe, en s'appuyant sur les différentes intensités observées, comme pour les méthodes SIR et MIR. En SAD, on collecte les données à une seule longueur d'onde, au maximum d'absorption du diffuseur anormal, tandis qu'en MAD on collecte aussi les données au point d'inflexion et à une troisième longueur d'onde éloignée du pic. L'enregistrement d'un spectre de fluorescence proche du seuil d'absorption permet de choisir les longueurs d'onde à appliquer. Enfin, d'autres méthodes combinées sont possibles mais ne seront pas détaillées, c'est le cas des méthodes SIRAS et MIRAS. Une fois que les phases sont déterminées, une première carte de densité électronique est calculée.

6.2.4 L'affinement

L'étape d'affinement permet ensuite d'améliorer le modèle global en suivant un processus itératif durant lequel les positions des atomes dans le modèle sont ajustées jusqu'à ce que les

amplitudes du facteur de structures calculées à partir du modèle s'accordent au mieux avec celles mesurées. La correspondance entre le modèle et les données expérimentales est suivie tout au long de l'affinement grâce au facteur d'accord cristallographique R. Plus R est petit plus l'accord est bon. Pour s'assurer de la qualité de l'affinement, on utilise aussi le facteur R_{free} , qui est calculé à partir de 5% de réflexions qui sont omises lors de l'affinement. Pour un modèle de bonne qualité, R_{free} (toujours supérieur à R) doit diminuer conjointement à R.

6.3 Cristallisation et collecte des données

Pour les tests de cristallisation de rPhoSL, rPhoNLSeMet, rPhoSLm et rPhoNLm, les protéines ont été concentrées entre 5.5 et 6.3 mg.mL⁻¹ dans 25 mM MES pH6 et pré-incubées, ou non, avec 1 mM de N-glycane fucosylé en α 1,6 (**1, 2, 3, 4, 5 ou 6, voir Figure III-8**). La technique de diffusion de vapeur en goutte suspendue a été utilisée en réalisant des gouttes de 2 μ l contenant 50 : 50 (v/v) d'un mélange de protéines et de solution réservoir à 19°C. Les premiers criblages ont été réalisés avec des kits commerciaux (Clear Strategy Screen I et II et Structure Screen 2, Molecular Dimensions Ltd) sans succès. Le robot de cristallisation (HTXLab, EMBL) a alors été utilisé pour un criblage de 576 conditions commerciales en réalisant des gouttes assises de 200 nl. Certaines conditions positives ont ensuite été reproduites au laboratoire et optimisées manuellement. Des cristaux de forme triangulaire ont été obtenus avec rPhoSL avec 18% butane-1,4-diol, 300 mM acétate de zinc et 0,1 M Imidazole-HCl pH 6-7, et des cristaux aux formes de diamants ont été obtenus avec différents complexes rPhoSL-ligands dans 3 à 3.4 M de malonate de sodium pH 5. Des cristaux sous forme de plaques ont été obtenus avec rPhoNLSeMet en complexe avec le N-glycane 1 et avec rPhoNLm-ligands dans 100 mM MgCl₂, 5% PEG 8K, 5% PEG 1K et dans 12.5 % PEG 2K, 200 mM MgCl₂, 0.1 M Tris pH 7 respectivement. Les cristaux ont ensuite été directement montés dans des boucles et congelés dans l'azote liquide car les conditions de cristallisation étaient cryoprotectantes.

Les données de rPhoSL ont été collectées sur la ligne BM14 à l'ESRF (European Synchrotron Radiation Facility), Grenoble, en utilisant un détecteur MARCCD. Les données de rPhoSL en complexe avec le N-glycane 1 ont été collectées sur la ligne ID23-1 de l'ESRF avec un détecteur pilatus 6M (Dectris Ltd), les données de rPhoSL en complexe avec les N-glycanes 2 et 3 et de rPhoNLm en complexe avec le N-glycane 6 ont été collectées sur la ligne Proxima-1 à SOLEIL, Saint Aubin sur un détecteur pilatus 6M (Dectris Ltd), et les données de rPhoSL en complexe avec le N-glycane 5 ont été collectées sur la ligne FIP-BM30 de l'ESRF avec un détecteur ADSC Q315r CCD. Les données de rPhoNLSeMet ont été

collectées sur la ligne FIP-BM30 de l'ESRF avec un détecteur ADSC Q315r CCD. Les clichés de diffraction ont été traités avec XDS (Kabsch, 2010) et tous les calculs suivants ont été réalisés avec la suite de logiciels CCP4i (Winn et al., 2011).

6.4 Résolution des structures et affinement

Les structures de rPhoSL et de rPhoNLSeMet ont été résolues par la méthode SAD avec le logiciel ShelXC/D (Sheldrick, 2010). Le signal anomal du zinc a été utilisé pour la structure de rPhoSL et le signal anomal du sélénium a été utilisé pour résoudre la structure de rPhoNLSeMet. Un modèle polyalanine a d'abord été construit par ShelXE et a ensuite été amélioré avec un assignement de la séquence en utilisant ArpWarp (Langer et al., 2008). La technique de remplacement moléculaire a été utilisée pour résoudre les structures de rPhoSL et rPhoNLm en complexe avec les différents N-glycanes fucosylés en α 1,6 à l'aide du logiciel PHASER (McCoy et al., 2007). L'affinement des structures a été réalisé avec Refmac5.8 (Murshudov et al., 2011) et les corrections et constructions manuelles des modèles ont été réalisées dans COOT (Emsley et al., 2010). Les sucres ont été ajoutés manuellement et les molécules d'eau de façon automatique dans COOT. La qualité stéréochimique des modèles affinés a été validée sur le serveur wwPDB Validation : <http://wwpdb-validation.wwpdb.org> et les sucres ont été validés dans Privateer (Agirre et al., 2015). Toutes les figures ont été dessinées dans PyMOL Molecular Graphic System program (version 2.0.4, Schrodinger, LLC).

7. Mesure de l'activité anti-adhésive pour des micro-organismes

L'activité anti-adhésive de rPhoSL a été testée sur onze micro-organismes différents (*E. coli*, *B. cereus*, *B. subtilis*, *Campylobacter*, *Salmonella*, *Pseudomonas*, *L.monocytogenes*, *A. flavus*, *A. ochraceus*, *P. anomala*, *Candidat intermedia*) avec rPhoSL à 200, 20 et 2 $\mu\text{g.mL}^{-1}$. rPhoSL a été incubée avec les micro-organismes dans des plaques 96 puits avec un contrôle positif sans lectine et un contrôle négatif sans micro-organisme. Des dilutions en séries au centième ont été réalisées pour chaque condition et les plaques ont été incubées pendant 24h à la température optimale de pousse du micro-organisme testé. Les puits ont alors été lavés trois fois avec une solution saline puis passés au sonicateur 10 min. Une nouvelle dilution en série au dixième de chaque condition a ensuite été réalisée, puis 10 μl par condition ont été déposés sur boîte de pétri et placés dans un incubateur pendant 24 h. Les colonies ont ensuite été comptées. La diminution du nombre de colonies par rapport au contrôle positif indique une activité anti-adhésive de rPhoSL.

8. Étude biologique sur cellules et tissus cancéreux

8.1 Cytométrie en flux

Les lectines rPhoS_LM, rPhoN_LM et rTrfbL2 (WT et Y111A) ont été testées en cytométrie en flux afin d'évaluer leur capacité de reconnaissance spécifique de leur ligand sur des cellules cancéreuses.

8.1.1 Principe de la cytométrie en flux

La cytométrie de flux est une technique qui permet de mesurer dans un mélange de cellules, les caractéristiques individuelles de chacune d'entre elles telles que la taille, la granulométrie, ou la présence d'un ligand si celui-ci peut être détecté par une molécule fluorescente.

Les cellules à étudier sont mises en suspension dans un tampon adéquat et poussées les unes après les autres devant un faisceau laser. Des capteurs mesurent la lumière diffractée aux petits angles (Forward Scatter FSC) et aux grands angles (Side Scatter SSC). Le paramètre FSC mesure la taille des cellules et le paramètre SSC mesure la granularité des cellules (leur complexité) ainsi que la fluorescence et/ou l'autofluorescence. L'appareil de FACS est alors capable de trier les cellules selon les différentes caractéristiques détectées directement à partir de critères choisis par l'utilisateur. Pour cela, la veine liquide est chargée électriquement et une onde de vibration fractionne le flux en une succession de gouttelettes (une cellule par gouttelette). Les cellules d'intérêt à trier passent ensuite entre des plaques fortement chargées et sont déviées du côté de la plaque de polarité opposée avant d'être collectées. Pour les cellules non sélectionnées, la veine liquide n'est pas chargée et elles seront donc éliminées (*Figure III-12*). Le mode de tri peut être modifié pour améliorer la pureté ou le rendement. Dans notre cas, les cellules ont été incubées avec les lectines marquées avec un fluorophore avant d'être analysées. Le taux de fluorescence des cellules mesuré est proportionnel au nombre de lectines fixées et donc au nombre de ligands accessibles en surface.

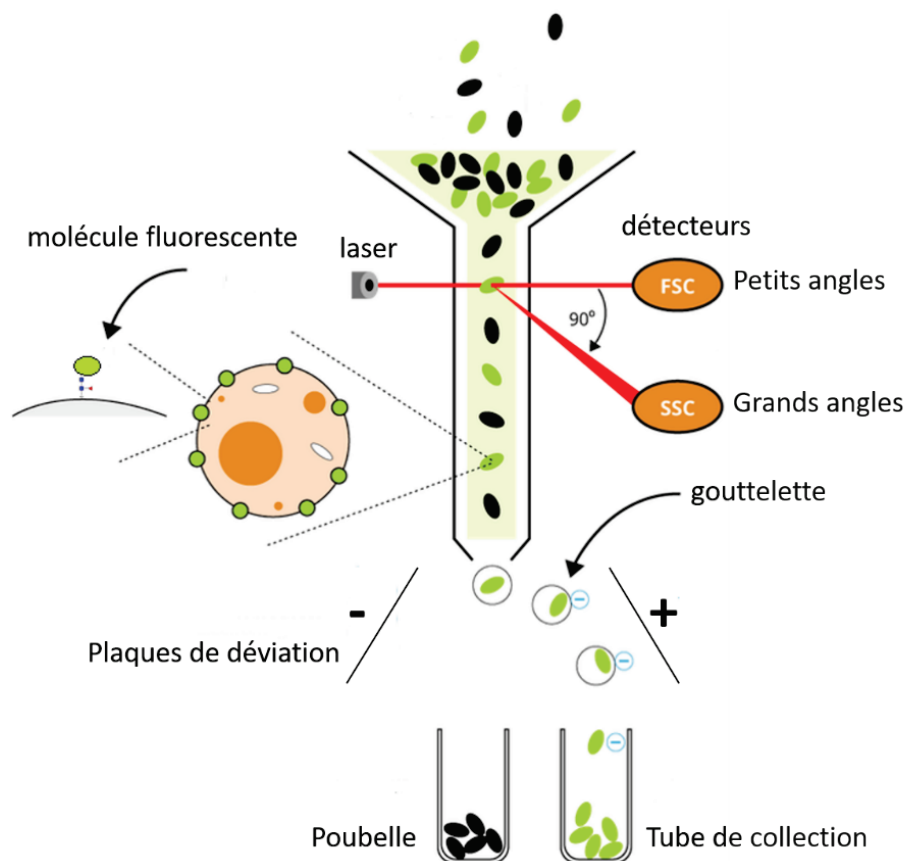


Figure III-12. Principe du FACS.

8.1.2 Traitement des échantillons

Différentes lignées cellulaires ont été utilisées : Les lignées MINO et Z-138 dérivées du lymphome du manteau (MCL), HT-29 dérivées de l'adénocarcinome du côlon, MCF7, dérivées de tumeurs mammaires, et A549, issue de cellules épithéliales pulmonaires adénocarcinomiques. Les différents milieux de cultures utilisés sont décrits **Tableau III-6** pour chaque lignée. Toutes les lignées cellulaires ont été incubées à 37°C avec 5% de CO₂ dans un incubateur humidifié.

Lignée cellulaire	Milieu de culture
MINO	RPMI1640 + 10% FBS, 1% Pénicilline / Streptomycine et 1% L-Glutamine
Z-138	
HT-29	DMEM à 4,5 g.L ⁻¹ de glucose + 10% FBS, 1% de Pénicilline / Streptomycine et 1% L-glutamine
MCF7	DMEM-F12-Glutamax + 10% FBS
A549	RPMI-Glutamax + 10% FBS

Tableau III-6. Milieux de cultures utilisés pour les différentes lignées cellulaires utilisées en cytométrie en flux

Des mesures ont été réalisées à l'IAB (Institute for Advanced Biosciences, Grenoble) où les cellules ont été centrifugées à 60-80% de confluence, lavées brièvement dans du PBS 1X et resuspendues dans de la trypsine. Elles ont ensuite été centrifugées à 300 g pendant 5 min et lavées avec du PBS 1X. 5×10^5 cellules par tube ont alors été suspendues dans 100 μ l de différentes concentrations de lectines marquées à l'Alexa 680 dans du PBS 1X et incubées durant 30 min à 37°C. Après un double lavage, elles ont été suspendues dans 200 μ l de PBS et leur intensité de fluorescence a été mesurée à l'aide d'un cytomètre de flux BD C6 Accuri.

Des expériences d'inhibition de l'expression de la fucosyltransférase 8 ont aussi été réalisées. Pour cela, des siRNA (ON-TARGETplus Human FUT8 (2530) siRNA-SMART pool, Dharmacon) ont été transfectés dans les cellules cibles. Ils dirigent le clivage des ARNs de la FUT 8, empêchant la synthèse de l'enzyme. L'expérience a été réalisée sur les cellules MCF7 et A549. Les cellules ont été transfectées à 30-50 % de confluence avec 50 nM de siRNA FUT8 ou de siRNA LUC (luciférase, contrôle positif de la transfection) par condition, en utilisant du jetPRIME (Poluplus) comme agent de transfection. Les cellules ont été incubées avec les siRNA à 37°C pendant 72 h avant d'être récupérées et incubées avec les lectines.

D'autres mesures ont été réalisées au CRCINA (Centre de Recherche en Cancérologie et Immunologie Nantes-Angers, Nantes) où les cellules ont été comptées puis centrifugées à 300 g pendant 5 min et remises en suspension dans du tampon PBS-0.1% BSA à une concentration de 2.5×10^6 cellule/ mL. Elles ont alors été réparties dans des plaques 96 puits à raison de 100 μ l/puits. Les cellules ont ensuite été incubées avec les lectines biotinyllées à 0.1 ou 0.2 μ g.mL⁻¹ dans du PBS-0.1 % BSA pendant 30 min à 1 heure à 4°C. Après trois lavages avec du PBS-0.1 % BSA, les cellules ont été incubées avec de la streptavidin-PE (streptavidin-phycoerythrin ; 0,5 mg.mL⁻¹; BD Bioscience) au 1/500^{ème} dans du PBS-0.1 % BSA pendant 30 min à 4°C et enfin lavées trois fois dans du PBS. L'analyse de cytométrie de flux a été réalisée sur un appareil de cytométrie de flux BD C6 Accuri.

Des expériences sur PBMC (cellules mononuclées du sang périphérique) ont également été mises en place au CRCINA. Les PBMC ont été préparés à partir de sang frais collecté chez des donneurs sains (Établissement Français du Sang, France) et accompagné d'un consentement écrit. Les cellules ont été isolées par gradient de densité avec du Ficoll (Eurobio) puis marquées avec un mélange contenant 1/100 d'anti-CD36PE-cyanin7 (BD bioscience), 1/20 d'anti-CD19-BV605 (BD bioscience), 1/20 d'anti-CD14-FITC (BD

bioscience) et rPhoSL-Alexa680 à 20 $\mu\text{g}\cdot\text{mL}^{-1}$ dans du PBS-0.1% BSA. L'analyse a été réalisée avec un cytomètre de flux BD LSRIL.

8.2 Lectine-histochimie

Les lectines rPhoSL, rPhoSLm, rPhoNLm, rTrfbL2-WT et rTrfbL2-Y111A ont été testées en lectine-histochimie pour leur capacité de reconnaissance spécifique de leur ligand sur tissus cancéreux.

Des premières mesures ont été réalisées au CRCINA de Nantes avec des échantillons humains de trachée, œsophage, jonction duodénale, jéjunum, côlon, pancréas, foie, ovaire, utérus et vagin, obtenus à partir de donneurs d'organe avant la loi n°88-138 du 20 décembre 1988 concernant la résection de tissus humains après le décès à des fins d'investigation scientifique. Ils ont été utilisés pour préparer des « tissue microarray » (TMA). 18 échantillons de tissus de 10 individus différents ont été utilisés. Des coupes de tissus tumoraux issus du côlon ont aussi été testées. Les sections ont été déparaffinées dans des bains de LMR Sol (Labo-moderne, Paris), puis dans des bains de solutions d'éthanol à des pourcentages décroissants (100, 90, 70 et 50 %). Les fixations non-spécifiques ont été bloquées avec de la BSA à 5 % dans du PBS pendant 60 min. rPhoNLm diluée à 20 $\mu\text{g}\cdot\text{mL}^{-1}$ dans du PBS-1% BSA a été ajoutée et incubée 60 min à température ambiante. En condition d'inhibition, rPhoNLm a été pré-incubée avec 25 mM de chitobiose fucosylé en $\alpha 1,6$, pendant 60 min à 4°C afin de bloquer ses sites de fixation avant d'être ajoutée sur les lames. Après trois lavages avec du PBS, de la streptavidine conjuguée à l'Alexa549 (1 $\text{mg}\cdot\text{mL}^{-1}$, Vector Laboratories, Burlingame, CA) diluée au 1/200^{ème} dans du PBS-1% BSA, a été ajoutée pendant 60 min à température ambiante. Après trois autres lavages avec du PBS, les lames ont été montées dans du ProLong Gold (Thermo). Elles ont été observées sur un microscope Zeiss « ApoTome » ou un scanner Nanozoomer (Hamamatsu, Hamamatsu City).

Des expériences complémentaires ont ensuite été réalisées à l'IAB sur des coupes de tissus cancéreux. Ces coupes sont issues de tumeurs développées chez le rat suite à des injections de cellules de lignées tumorales humaines : A549, HT29, MCF7 et U87. Des tests sur des coupes de foie sain murin ont aussi été testées. Les coupes ont préalablement été décongelées puis fixées à l'acétone. Après trois rinçages dans du TBST 1X (Tris buffer saline-Tween), les lames ont été cerclées puis l'activité des peroxydases endogènes a été bloquée avec du BLOXALL (sp-6000 VectorLabs) pendant 10 min. Les coupes ont ensuite été rincées avec du TBST 5 min avant d'ajouter du sérum de cheval (2.5%, VectorLabs S-2012) et laissées

incuber 10 min. Lorsque le blocage des biotines endogènes était nécessaire, de la streptavidine à 0.1 mg.mL^{-1} puis de la biotine à 0.5 mg.mL^{-1} ont été ajoutées pendant 15 min chacune. Les lectines rPhoS_Lm-peroxydase, rPhoN_Lm-biotinylée, rTrfbL2-WT-peroxydase et rTrfbL2-Y111A-peroxydase à $20 \text{ } \mu\text{g.mL}^{-1}$ ont ensuite été incubées avec les coupes sur la nuit à 4°C . Lors des expériences d'inhibition, elles ont été pré-incubées pendant minimum une heure avec leur ligand (25 mM de chitobiose fucosylé en $\alpha 1,6$, 25 mM d'antigène TF, ou 100 mM GlcNAc) avant d'être déposées sur les coupes. Les coupes ont ensuite été lavées au TBST avant l'étape de révélation. Dans le cas de rPhoN_Lm-biotinylée, de la streptavidine-peroxydase a été ajoutée au $1/250^{\text{ème}}$ 30 min à température ambiante, puis, pour toutes les lectines, les coupes ont été révélées avec de l'immPACTNovaRED (VectorLabs SK-4805) selon les instructions du fournisseur. Les coupes ont alors été rincées à l'eau 5 min avant d'ajouter de l'hématoxyline pendant une minute afin de colorer les noyaux. Après deux rinçages de trois min dans l'éthanol 100 %, elles ont été lavées au xylène pendant 15 min. Les lames ont ensuite été préparées au Pertex puis observées avec un scanner ZEISS Axioscan Z1.

IV.

RÉSULTATS ET

DISCUSSION

Chapitre 1.

**Outil de reconnaissance de
l'antigène TF : rTrfbL2**

1. Outil de reconnaissance de l'antigène TF : rTrfbL2

TrfbL2 est une lectine qui présente un fort potentiel dans la reconnaissance spécifique de l'antigène TF, un marqueur cancéreux. L'objectif de ce travail était de l'étudier, l'analyser et la modifier afin de créer un outil performant, très spécifique et facile à produire.

1.1 Caractéristiques de TrfbL2 et production sous forme recombinante

La protéine TrfbL2 a tout d'abord été extraite du champignon *Terfezia boudieri* (TrfbL2 native) pour être caractérisée et sa structure a été résolue par cristallographie aux rayons X. Afin d'assurer sa qualité, sa quantité et un approvisionnement régulier, elle a ensuite été produite sous forme recombinante (rTrfbL2) dans *E. coli*. Ce mode de production est très efficace puisqu'après purification par chromatographie d'affinité sur une résine à GlcNAc, le rendement est d'environ 20 mg par litre de culture. La forme recombinante de la lectine présente les mêmes caractéristiques que TrfbL2 native comme la formation de tétramères, une grande stabilité avec une température de fusion à laquelle 50% des protéines sont dénaturées (T_m) proche de 80°C, et une double spécificité pour le GlcNAc et le GalNAc. Une étude calorimétrique (ITC) a montré que rTrfbL2 avait une affinité de l'ordre du micromolaire pour l'antigène TF (Gal β 1-3GalNAc, code GlyTouCan G58645HA). Elle a besoin du disaccharide pour la reconnaissance puisqu'elle ne reconnaît pas le GalNAc seul ni même l'antigène Tn (GalNAc α 1 O-Ser/Thr). Il est en effet probable que les interactions entre le GalNAc et TrfbL2 ne soient pas suffisantes pour stabiliser le monosaccharide et que les interactions supplémentaires avec le galactose soient nécessaires à l'interaction. L'affinité pour le GlcNAc est seulement de l'ordre du millimolaire (800 μ M mesurés par ITC pour du GlcNAc seul et 500 μ M pour du GlcNAc β 1-3Gal).

TrfbL2 présente des homologies de séquence et de structure avec les lectines de champignon BEL de *Boletus edulis* (Bovi et al., 2011), ABL d'*Agaricus bisporus* (Carrizo et al., 2005), XCL de *Xercomus chrysenteron* (Birck et al., 2004) et SRL de *Sclerotium rolfsii* (Leonidas et al., 2007) (**Figure IV-1**). Ces lectines ont été décrites comme formant des dimères de dimères et sont caractérisées par la présence de deux sites de fixation distincts par protomère, le premier étant spécifique du GalNAc et le second spécifique du GlcNAc. Ces lectines appartiennent à une même famille avec un repliement de type actinoporine. Ce repliement est caractérisé par la formation d'un β -sandwich dont les deux feuilletts β sont connectés entre eux via un motif hélice-tour-hélice.

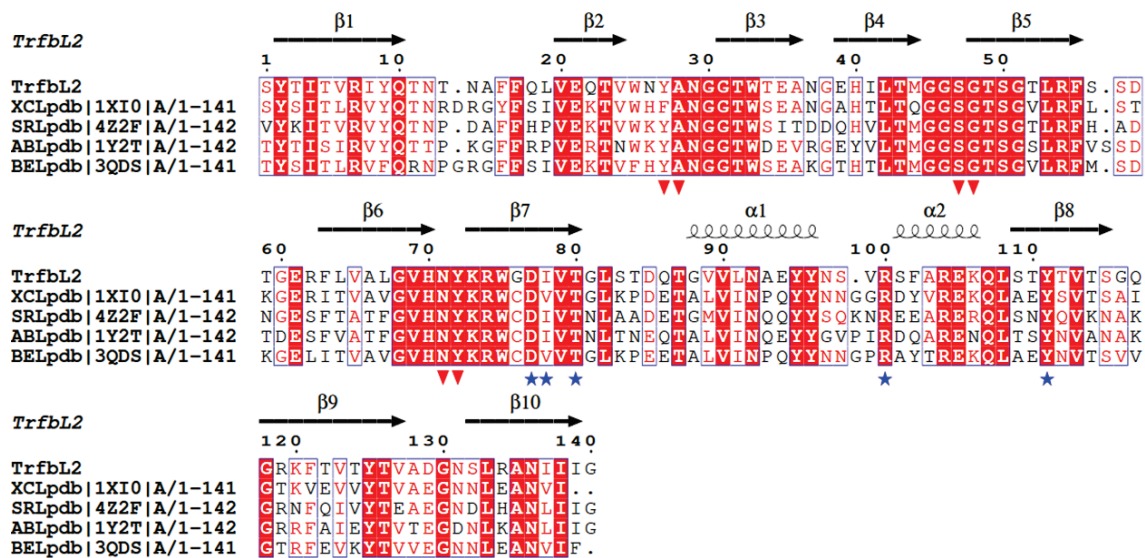


Figure IV-1. Alignement des séquences homologues à TrfbL2 de structure connue. Alignement réalisé avec Esript3 (Robert and Gouet, 2014). Numérotation de TrfbL2 et structures secondaires au dessus. Résidus en interaction avec l'antigène TF et le GlcNAc indiqués par un triangle rouge et une étoile bleue respectivement.

Le domaine β -sandwich de TrfbL2 est constitué d'un premier feuillet composé des 6 brins H (résidus 109-115), I (résidus 118-127), J (résidus 132-139), A (résidus 2-10), D (résidus 39-44) et C (résidus 31-36) et d'un deuxième feuillet composé des 4 brins B (résidus 20-24), E (résidus 48-55), F (résidus 63-70) et G (résidus 73-80). Un motif hélice-tour-hélice relie entre eux les deux feuillets β en joignant les brins G et H. L'espace entre les deux feuillets est principalement constitué de chaînes latérales hydrophobes (*Figure IV-2*).

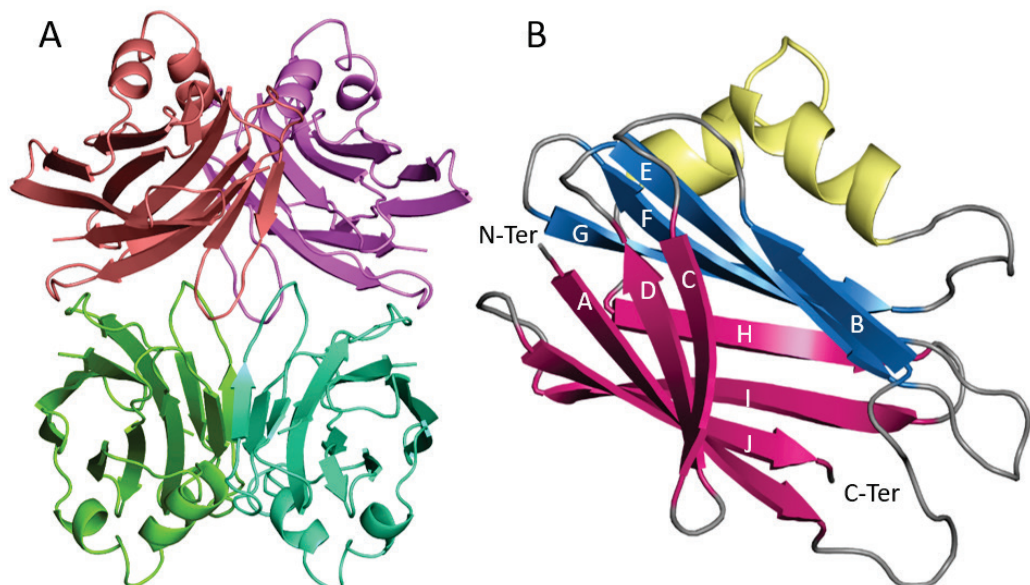


Figure IV-2. A. Structure du tétramère de TrfbL2 native. B. Description de la structure d'un monomère de TrfbL2 native.

A. Coloration par protomère. B. Le premier feuillet constitué de 6 brins est en rose, le second de 4 brins en bleu. Le motif hélice-tour-hélice en jaune.

Les deux sites de reconnaissance de TrfbL2 sont situés de part et d'autre du motif hélice-tour-hélice. Le site au GalNAc que l'on appellera site 1 forme un renforcement peu profond délimité par les boucles connectant les brins B→C, D→E et F→G. Le site de reconnaissance du GlcNAc que l'on appellera site 2 est délimité par les derniers résidus du brin G, le brin H et la deuxième hélice du motif hélice-tour-hélice (*Figure IV-3 A, D et E*).

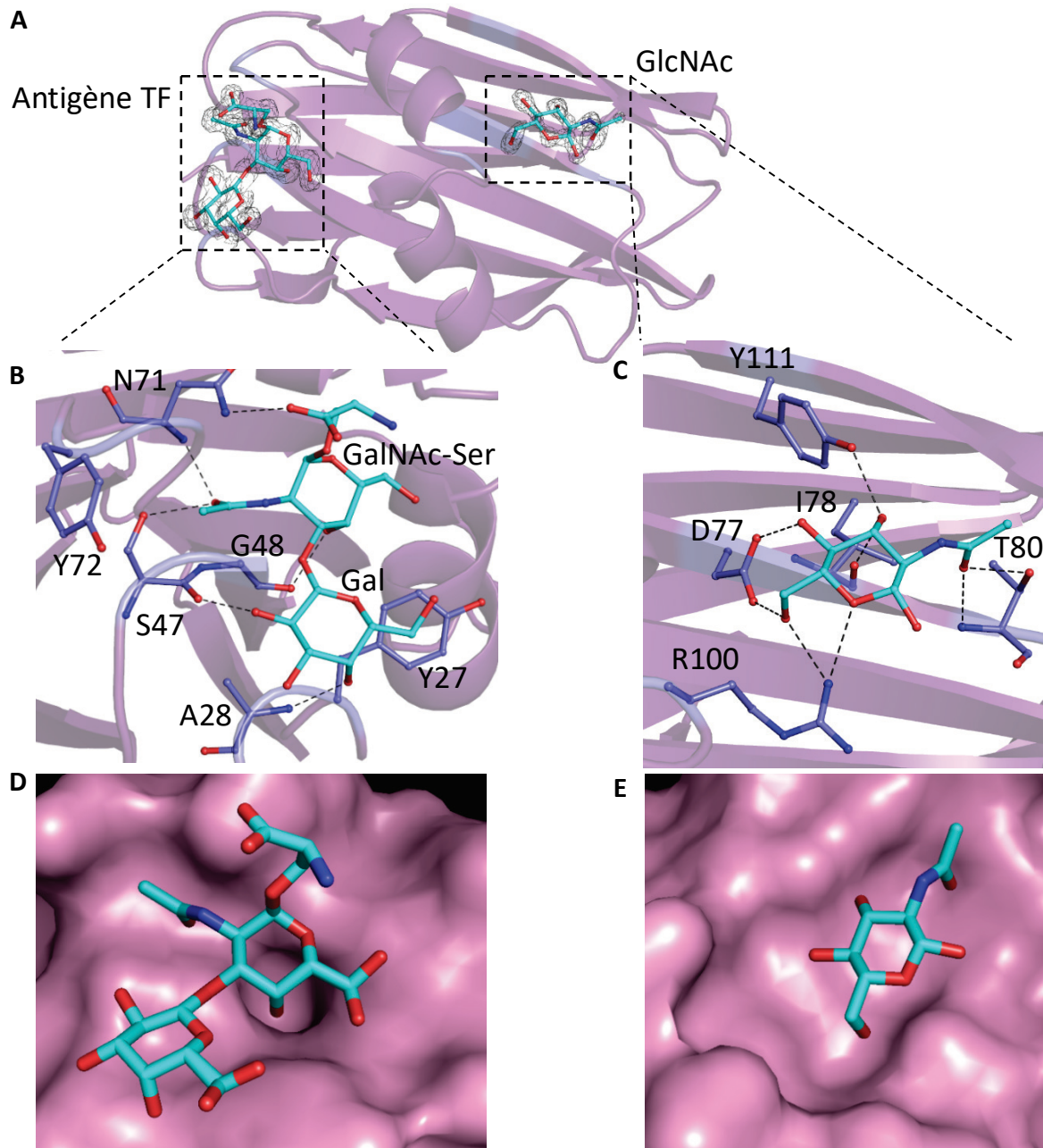


Figure IV-3. Description des interactions entre TrfbL2-WT et ses ligands.

A. Structure d'un monomère de TrfbL2-WT en présence de GlcNAc et de l'antigène TF lié à une sérine (TF-Ser) en densité. B. Description de l'interaction entre TrfbL2-WT et le TF-Ser. C. Description de l'interaction entre TrfbL2-WT et le GlcNAc. D. Représentation de surface du site d'interaction avec le TF-Ser. E. Représentation de surface du site d'interaction avec le GlcNAc.

Les interactions avec l'antigène TF et le GlcNAc sont les mêmes dans toutes les structures homologues contenant ces sucres (TrfbL2, BEL et ABL). Sur le site 1, le GalNAc de l'antigène TF interagit avec le OG de la sérine 47 et le N de l'asparagine 71 via son O7, et avec le carbonyle de la glycine 48 via son O4. Le galactose de l'antigène TF interagit avec l'hydroxyle de la sérine 47 via son O2 et avec l'alanine 28 via son O4. Seuls deux résidus hydrophobes, la glycine 48 et l'alanine 28, sont présents au fond du site de fixation 1. On observe également un effet d'empilement entre le galactose et le cycle de la tyrosine 27. Au niveau du deuxième site d'interaction, le GlcNAc interagit avec les oxygènes de l'aspartate 77 via son O4 et son O6, avec l'hydroxyle de la tyrosine 111 via son O3, avec le OG1 et l'azote de la thréonine 80 via son O7 et avec le NH1 de l'arginine 100 via son O5. Les résidus hydrophobes présents sur le site 2 : la valine 79, l'isoleucine 78, la valine 113 et l'alanine 103, participent à la formation d'une poche d'interaction peu profonde (**Figure IV-3 B, C**).

TrfbL2 et ses homologues sont caractérisées par leur remarquable spécificité pour le GalNAc sur le site 1 et pour le GlcNAc sur le site 2. Le GalNAc et le GlcNAc sont en effet des sucres très similaires dont la seule différence réside dans l'orientation de l'hydroxyle en position 4 qui est équatoriale pour le GlcNAc et axiale pour le GalNAc (**Figure IV-4**). Du fait de la différence d'orientation de cet hydroxyle, la fixation d'un GalNAc à la place du GlcNAc dans le site 2 n'est pas possible principalement car le O4 entrerait en conflit stérique avec l'aspartate 77. Selon le même principe, la fixation d'un GlcNAc au lieu du GalNAc dans le site 1 est impossible à cause d'un conflit stérique avec la tyrosine 27.

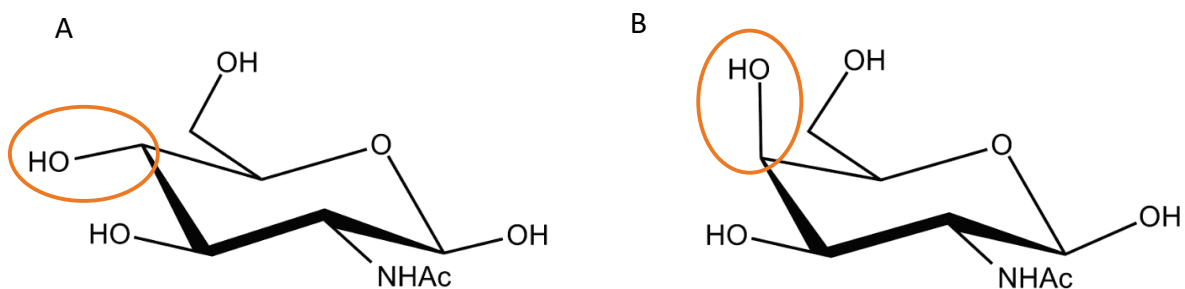


Figure IV-4. Structure du GlcNAc (A) et du GalNAc (B).

rTrfbL2 a aussi été cristallisée en interaction avec du GlcNAc β 1-3Gal (**Figure IV-5.A**). Ce sucre interagit sur le site 2 et la structure montre que le O4 du galactose fait des interactions supplémentaires avec rTrfbL2-WT via l'arginine 100. Ce disaccharide est présent dans les structures de N-glycanes que l'on retrouve en surface des cellules cancéreuses (**Figure IV-5.B**), tout comme les GlcNAc terminaux (Audfray et al., 2015).

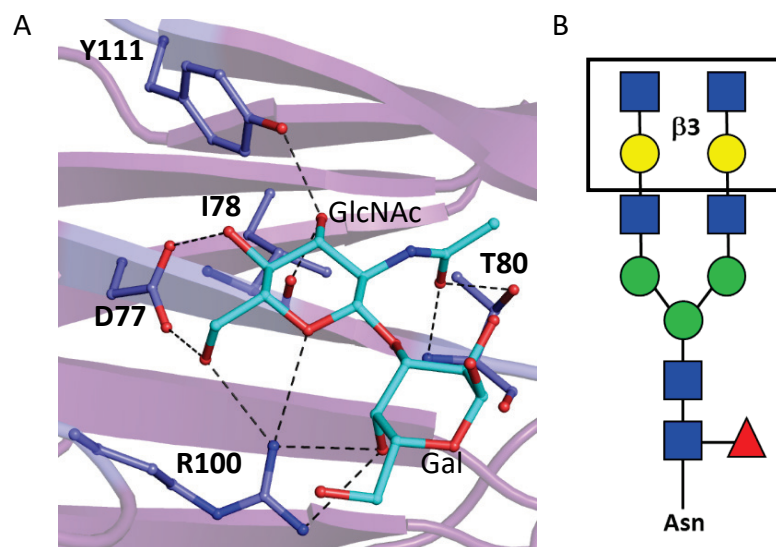


Figure IV-5. A. Interaction de rTrfbL2-WT avec du GlcNAc β 1-3Gal. B. Schéma d'un N-glycane possédant le disaccharide GlcNAc β 1-3Gal.

Étant donné que TrfbL2 reconnaît des motifs glucidiques retrouvés à des taux anormaux dans les cancers (antigène TF et GlcNAc terminaux), elle présente un grand intérêt comme outil biotechnologique en glycosciences surtout pour la reconnaissance spécifique de l'antigène TF. La spécificité secondaire pour le GlcNAc présente moins d'intérêt mais pourrait gêner et biaiser la reconnaissance de l'antigène TF. Il est donc recommandé d'invalider cette spécificité secondaire autant que possible afin d'avoir un outil des plus spécifiques.

1.2 Conception de mutants abrogeant la reconnaissance du GlcNAc

La stratégie adoptée pour invalider le site 2 de reconnaissance du GlcNAc a été de réaliser des mutations ponctuelles sur des acides aminés stratégiques, en s'appuyant sur l'analyse des séquences homologues et sur les structures de TrfbL2.

L'analyse des séquences a montré que tous les résidus en interaction avec le GlcNAc décrits *Figure IV-3* sont conservés chez les lectines homologues (*Figure IV-1*). Seuls les trois acides aminés D77, R100 et Y111 ont été désignés pour les mutations, car les deux autres acides aminés en interaction (T80 et I78) n'interagissent pas via leur chaîne latérale. Les conséquences des mutations sur la stabilité de la protéine n'étant pas connues, un panel de mutants a été créé : Y111Q, Y111A, D77A, Y111QD77A, et R100G. Ils ont été conçus par mutagenèse dirigée à partir du plasmide de rTrfbL2-pET29a puis clonés ou non dans différents vecteurs d'expression (Pet-TEV, pET32_TEV, pProEx-HtB). La description des vecteurs utilisés est située dans le *chapitre III*.

1.2.1 Production et purification de rTrfbL2-Y111A et rTrfbL2-Y111Q

La mise au point des protocoles de production et de purification des protéines rTrfbL2-Y111A et rTrfbL2-Y111Q a été réalisée en se basant sur le protocole établi pour rTrfbL2 non mutée. Les mutants ont été exprimés sous forme recombinante dans la souche bactérienne Rosetta pLysS (DE3) *d'E. coli*, puis purifiés par chromatographie d'affinité sur une résine à GlcNAc, un des ligands reconnu par rTrfbL2. Ces deux mutants sont toujours capables d'héماغglutiner les globules rouges, comme la protéine sauvage, et leur stabilité n'est pas altérée selon les mesures par TSA.

Des mesures d'ITC ont aussi permis de vérifier que ces mutants étaient toujours capables de reconnaître l'antigène TF avec la même affinité que la protéine sauvage. Des essais de mesures d'affinité pour le GlcNAc ont été réalisés en ITC avec rTrfbL2-Y111Q et rTrfbL2-Y111A, sans succès, ce qui montre que les mutations permettent de diminuer l'affinité de ces protéines pour le GlcNAc comparé à la protéine sauvage dans une gamme non détectable par ITC. Cependant les mutants Y111A et Y111Q peuvent toujours reconnaître le GlcNAc puisque l'on peut les purifier sur une colonne d'affinité au GlcNAc. Les mutants Y111Q et Y111A présentant des caractéristiques très similaires, la suite de l'étude ne sera décrite que pour le mutant Y111A.

1.2.2 Étude structurale du mutant Y111A

Afin de vérifier que la mutation Y111A n'altérerait pas la structure globale de la protéine, et pour mieux comprendre les interactions qui maintiennent la reconnaissance du GlcNAc, la lectine rTrfbL2-Y111A a été co-cristallisée avec du GlcNAc et de l'antigène TF à 1 mM. La co-cristallisation a été réalisée dans 25 % Peg 550 MME, 5 % Peg 20K, 50 mM AmSO₄ et 0.1 M Mes pH 6,5. Des cristaux en forme de pyramide ont été obtenus en quelques jours. Les données ont été collectées sur la ligne FIP-BM30 de l'ESRF avec un détecteur ADSC Q315r CCD. Elles appartiennent au groupe d'espace P4₁2₁2 et la structure a été résolue par remplacement moléculaire à une résolution de 1.7 Å.

Les données de collecte et d'affinement pour la structure du mutant Y111A sont décrites dans le **Tableau IV-1**. Il y a deux monomères dans l'unité asymétrique. L'affinement de la structure puis son analyse ont permis de confirmer que la mutation n'avait pas d'impact sur le repliement global de la protéine (rmsd de 0.634 entre rTrfbL2-WT et rTrfbL2-Y111A).

		rTrfbL2-Y111A	
Conditions de cristallisation		25% PEG 550 MME, 5% PEG 20K, 50mM AmSO ₄ , 0.1 M MES pH 6,5	
Données collectées			
Ligne de lumière		ESRF BM30A	
Groupe d'espace		P4 ₁ 2 ₁ 2	
Maille a b c (Å) α β γ (°)		78.67 78.67 118.22 90 90 90	
Longueur d'onde		1.008	
Resolution (Å)		39.32-1.7 (1.73-1.7)	
Réflexions mesurées		1189869	
Réflexions uniques		41461	
Rmerge		0.093 (0.554)	
Mean I / sd(I)		25.8 (4.1)	
Complétude (%)		99.7 (94.6)	
Multiplicité		28.7 (21.9)	
CC _{1/2}		0.999 (0.975)	
Affinement			
Résolution (Å)		37.32-1.7	
No. of work/ free Reflections		41545 / 2096	
R _{work} / R _{free} (%)		15.7 / 19.3	
Rmsd bond lengths (Å)		0.151	
Rmsd bond angles (°)		1.73	
Nb d'atomes		Chaîne A	Chaîne B
Protéine		1135	1096
Ligand		41	26
Autres hétéroatomes		20	
Eau		433	
B-facteurs moyens			
Protéine		20.18	18.46
Ligand		26.47	18.13
Autres hétéroatomes		40.59	
Eau		33.69	
Ramachandran plot (%)			
Allowed		100	
Favored		98.7	
Outliers		0	

Tableau IV-1. Données de collecte et d'affinement de la structure de rTrfbL2-Y111A.

Les valeurs entre parenthèses correspondent à la couronne de plus haute résolution. R_{free} a été calculé avec 5% des réflexions.

Les interactions avec l'antigène TF sont inchangées et les différences observées reposent essentiellement sur les interactions avec le GlcNAc. La densité électronique de rTrfbL2-Y111A n'a permis de modéliser qu'un seul GlcNAc en interaction avec l'un des deux monomères présents dans l'unité asymétrique, alors que ce sucre est bien présent en densité sur les quatre monomères de la protéine sauvage. Cela est en accord avec le fait que l'affinité

pour le GlcNAc après mutation est très faible puisque non mesurable par ITC. Après mutation, il n'y a plus d'interaction avec le résidu 111 modifié en alanine, ce qui prouve l'intérêt de cette mutation.

Les interactions avec les autres acides aminés D77, R100 et T80 sont quant à elles non modifiées (**Figure IV-6**). Elles ne sont pas suffisantes pour stabiliser le sucre dans tous les sites d'interactions au sein du cristal mais maintiennent la reconnaissance du GlcNAc lorsque celui-ci est présent en forte densité comme c'est le cas dans les résines de chromatographie. Cela est probablement dû à la multivalence de rTrfbL2 qui en formant des tétramères présente quatre sites d'interaction par molécule, ce qui crée un phénomène d'avidité sur la résine qui compenserait ainsi la faible affinité.

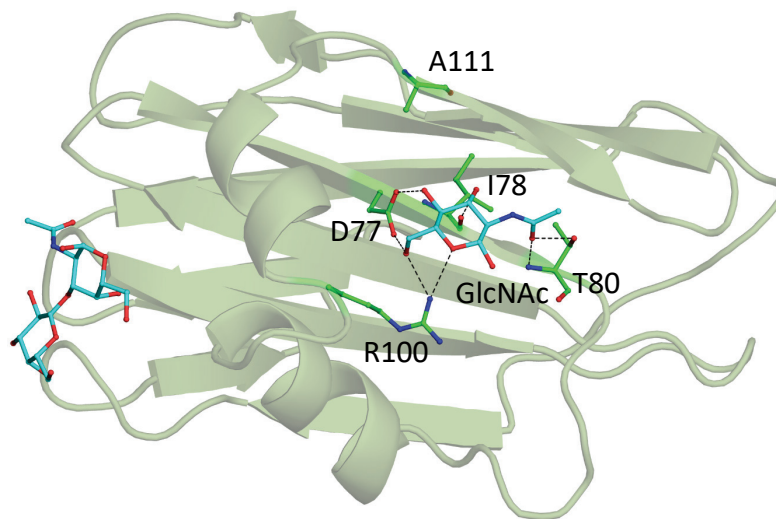


Figure IV-6. Interaction du GlcNAc avec rTrfbL2-Y111A (monomère).

La lectine a été co-cristallisée avec du GlcNAc et l'antigène TF. Seules les interactions avec le GlcNAc sont représentées.

1.2.3 Essais de purification de rTrfbL2-Y111QD77A et rTrfbL2- D77A

Afin de diminuer encore davantage l'affinité pour le GlcNAc, deux autres mutants ciblant cette fois l'aspartate 77 ont été analysés. Il s'agit des mutations Y111Q-D77A et D77A. Les deux mutants ne se fixent plus du tout sur la résine à GlcNAc utilisée auparavant pour purifier rTrfbL2, ce qui montre une perte d'affinité pour le GlcNAc plus importante que pour la mutation Y111A. Il est donc nécessaire de trouver d'autres méthodes efficaces de purification. Comme décrit dans le *chapitre III*, un grand nombre de techniques ont été testées. Ces mutants ont été d'abord clonés dans le vecteur pET29a, et produits dans les cellules Rosetta pLysS en raison de la présence de codons rares.

Des chromatographies d'affinité sur des résines Galactose, GalNAc ou sépharose ont été réalisées afin d'utiliser le site de fixation 1, sans succès. En supposant que les mutations n'ont pas altéré le repliement de rTrfbL2 et l'accessibilité du site 1, cela confirme la spécificité de rTrfbL2 pour l'antigène TF. La protéine sauvage ne se fixe pas non plus sur colonne galactose et sépharose, ce qui renforce cette conclusion.

Des tests ont ensuite été réalisés sur des colonnes de chromatographie échangeuses d'ions, pour essayer de purifier les lectines grâce à leurs charges. Malheureusement, aucune des colonnes échangeuses de cations ou d'anions testées n'a permis de séparer rTrfbL2 de l'ensemble des protéines d'*E. coli*. En effet, une faible quantité des mutants de rTrfbL2 se fixe sur les colonnes mais est éluée en même temps que les contaminants d'*E. coli*, soit parce que les charges sont similaires, soit à cause d'interactions non-spécifiques entre les mutants et les protéines d'*E. coli*. De plus, la majeure partie de rTrfbL2 mutée ne se fixe pas sur les colonnes échangeuses d'ions, quel que soit le pH du tampon et la colonne utilisée. rTrfbL2 a un point isoélectrique théorique égal à 8, ce qui signifie qu'elle est théoriquement chargée positivement à pH inférieur à 8, et qu'elle devrait se fixer aux colonnes échangeuses de cations à pH 5 à 7 comme cela a été testé. Afin de s'assurer que rTrfbL2 est stable à ces pHs, des mesures TSA ont été réalisées avec la protéine sauvage (**Figure IV-7**).

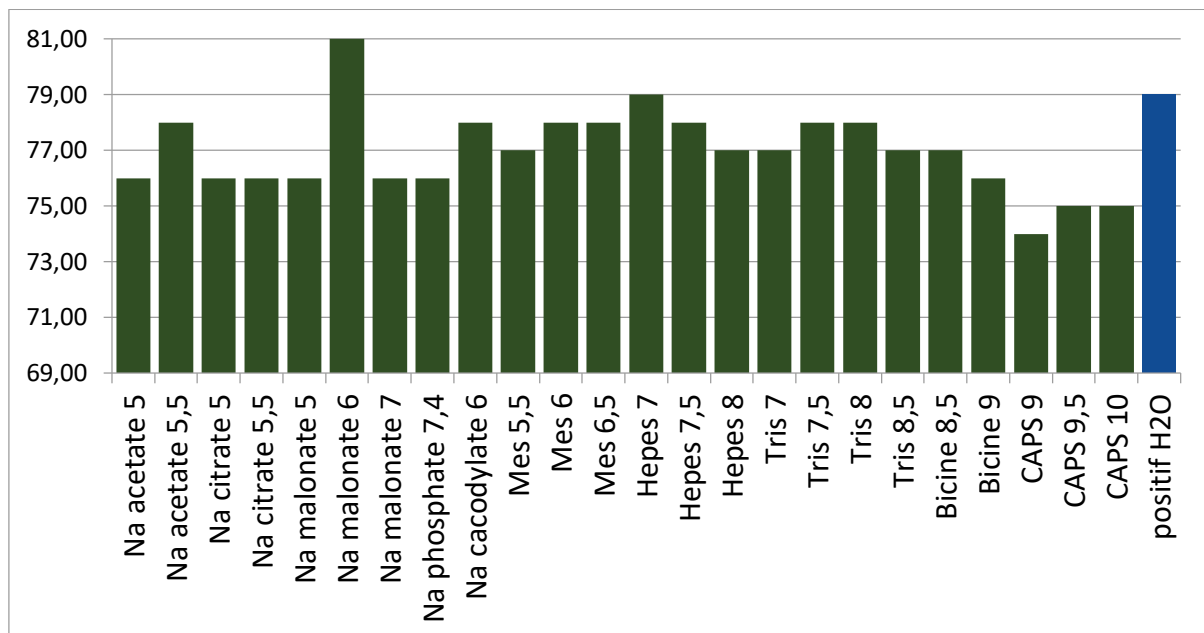


Figure IV-7. Thermal Shift Assay avec rTrfbL2-WT et une gamme de tampons de pH 5 à 10. Les tampons ont été dilués à 0.1 M. rTrfbL2 a été utilisée à 0.1 mg.mL⁻¹ dans du sypro orange 5x. Les mesures ont été réalisées sur un appareil de QRT-PCR (Mini-Opticon Biorad).

Les résultats montrent que rTrfbL2 a une température de fusion supérieure à 70°C pour tous les tampons entre pH 5 et pH 10. Le contrôle positif dans de l'eau (tampon de conservation de rTrfbL2) a une température de fusion de 79°C. Cela prouve la grande stabilité de rTrfbL2 à faible pH, et laisse supposer que le pH n'est pas en cause dans les difficultés des mutants de rTrfbL2 à se fixer aux colonnes échangeuses de cations. Pour confirmer cette théorie, la protéine rTrfbL2-Y111A pure a été passée sur une colonne échangeuse de cations CM avec un tampon 20 mM HEPES pH 7. Les résultats ont montré que l'ensemble de rTrfbL2-Y111A se fixait à la colonne, et que la protéine était éluée avec 300 mM de NaCl. rTrfbL2 est donc capable de se fixer aux colonnes échangeuses de cations en absence de contaminant. Les difficultés de fixation des mutants sur les colonnes de chromatographie échangeuses d'ions pourraient donc s'expliquer par des interactions protéine-protéine avec les contaminants, qui cacheraient les charges de rTrfbL2.

En supposant que les mutants rTrfbL2-Y111QD77A et rTrfbL2-D77A ont une grande stabilité thermique, comme la protéine sauvage, des tests ont été réalisés pour essayer de purifier ces mutants par la température. Le principe a été de faire chauffer les échantillons protéiques contenant rTrfbL2 mutée de 50 à 70°C pendant 15 min puis de centrifuger pour récupérer la fraction soluble. En supposant que la majorité des contaminants d'*E. coli* sont moins stables que rTrfbL2 mutée, rTrfbL2 mutée devrait rester dans la fraction soluble tandis que les contaminants d'*E. coli* précipiteraient. Cependant, les résultats ont montré qu'une grande partie des contaminants avaient la même stabilité thermique que les mutants rTrfbL2-Y111QD77A et rTrfbL2-D77A et qu'il n'était donc pas possible d'utiliser cette méthode pour les purifier. De plus, les mutants commencent à précipiter à 50°C et semblent donc moins stables que rTrfbL2 sauvage qui a une température de fusion de 80°C. La présence de précipités pendant les étapes de purification tend à confirmer cette hypothèse, même si les précipités peuvent correspondre aux autres protéines d'*E. coli*.

Toujours dans le but de purifier ces mutants et ainsi de pouvoir analyser leur stabilité en absence de contaminant, les mutants rTrfbL2-Y111QD77A et rTrfbL2-D77A ont été clonés dans les vecteurs pET-TEV, pProEX Htb et pET32-TEV qui permettent d'ajouter une étiquette poly-histidine en N-ter de la protéine. Cela permet ensuite de réaliser une purification sur colonne d'affinité au nickel. Le clonage a été réalisé dans le vecteur pET32-TEV car il permet une fusion à la thioredoxine qui aide au bon repliement des protéines. Les clones ont été transformés dans des cellules Rosetta pLysS et mis en culture selon le protocole décrit *chapitre III.2*. Après centrifugation des cultures, les cellules ont été lysées pour

récupérer l'extrait soluble qui a été appliqué ensuite sur une colonne d'affinité au nickel 1 ml cOmplete™ His-Tag (Roche). L'analyse des étapes de purification sur gel SDS PAGE a montré que rTrfbL2 mutée n'était pas ou très peu produite quel que soit le vecteur d'expression choisi, ce qui suggère que la mutation en elle-même, ou l'étiquette poly-histidine, freine le processus d'expression au sein de la bactérie. La souche Rosetta PlyS permet l'expression de protéines qui possède des codons rares, comme c'est le cas chez ces mutants, mais leur production n'est peut-être pas optimale dans ce type de souche. Pour pouvoir changer de souche d'expression et passer dans des cellules BL21(DE3) Star, le gène du mutant rTrfbL2-Y111QD77A a été à nouveau synthétisé avec une séquence optimisée qui ne possède plus de codons rares. Ce mutant a ensuite été cloné dans les trois vecteurs précédemment cités puis produit dans les cellules BL21(DE3) Star selon le même protocole avant de réaliser la purification sur une colonne d'affinité au nickel. Cette fois-ci, l'analyse sur gel SDS-PAGE a permis de repérer une bande éluee correspondant à la taille attendue de rTrfbL2 vers 33 kDa (15 kDa pour rTrfbL2 + 17.5 kDa pour l'étiquette en Nter) (**Figure IV-8**).

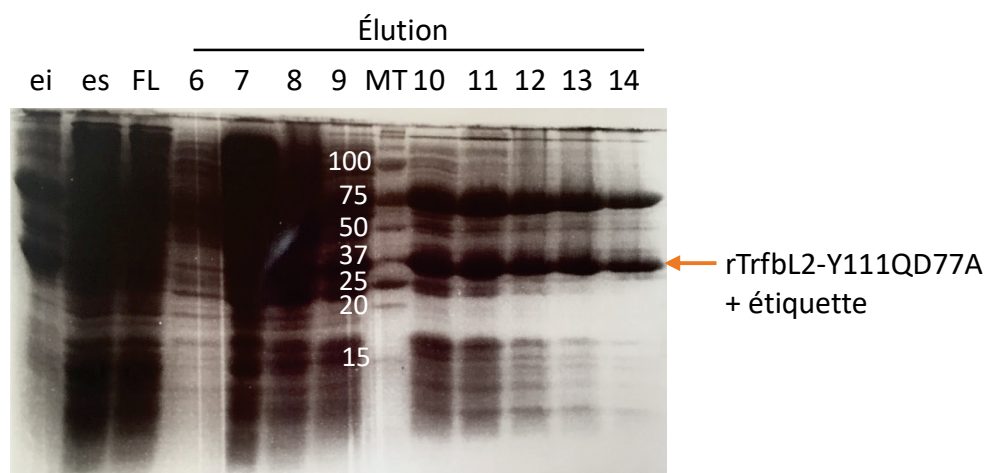


Figure IV-8. Gel SDS-PAGE Tris glycine 15% de la purification de rTrfbL2-Y111QD77A cloné dans le vecteur pET32-TEV sur Histrap.

ei = extrait insoluble ; es = extrait soluble ; FL = fraction non retenue ; MT = marqueur de taille en kDa.

L'étiquette a ensuite été coupée sur la nuit avec la protéase TEV et le mélange a été redéposé sur la colonne d'affinité au nickel. Après coupure et donc sans étiquette poly-histidine, rTrfbL2-Y111QD77A aurait dû ne plus être retenue par la résine et être collectée dans la fraction non retenue. Mais l'analyse sur gel SDS-PAGE a montré qu'aucune bande vers 15 kDa n'apparaissait dans cette fraction (**Figure IV-9**). En comparant les fractions coupées et non coupées, on remarque que deux bandes vers 13 kDa et 17 kDa s'intensifient

après coupure. La migration vers 13 kDa correspondrait au mutant rTrfbL2-Y111QD77A coupé : en effet rTrfbL2 migre toujours légèrement en dessous de sa taille, un phénomène observé couramment avec cette technique, et qui peut être dû aux charges de la lectine et ses interactions relatives avec le SDS. La bande vers 17 kDa correspondrait à l'étiquette seule. L'étiquette, qui possède une séquence poly-histidine, se fixe toujours sur colonne d'affinité au nickel et est donc retrouvée dans la fraction éluee. rTrfbL2-Y111QD77A coupée ne devrait plus avoir d'étiquette et ne devrait donc plus avoir d'affinité sur la colonne, elle est pourtant retrouvée elle aussi dans la fraction éluee uniquement. Cependant, rTrfbL2 formant des tétramères, si tous les monomères n'ont pas été coupés à la TEV, l'association d'un monomère coupé avec un monomère non coupé va permettre à la lectine de se fixer sur la colonne. On peut d'ailleurs remarquer qu'une bande correspondant à la taille de la protéine non coupée subsiste vers 33 kDa après coupure à la TEV.

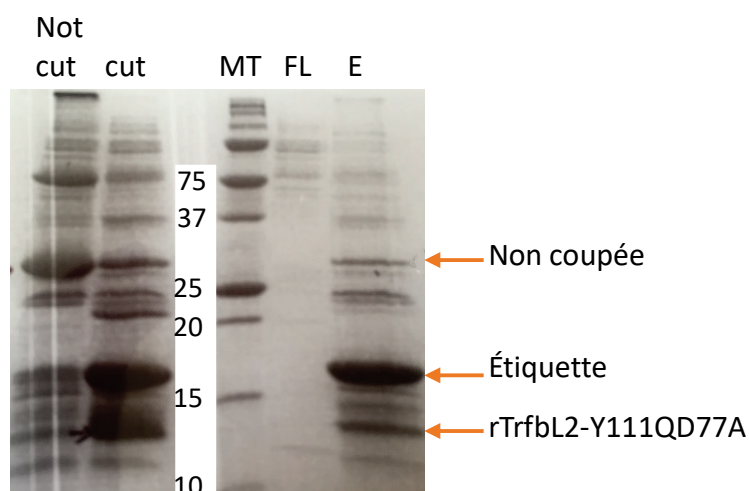


Figure IV-9. Gel SDS-PAGE Tris glycine 15% de la deuxième étape de purification de rTrfbL2-Y111QD77A cloné dans le vecteur pET32-TEV sur Histrap (après coupure de l'étiquette Nter).

Nc = non coupée ; cut = coupée ; MT = marqueur de taille en kDa; FL = fraction non retenue ; E = élution.

De plus, tout au long de la purification, les fractions sont très vite devenues troubles, signe de précipitation. Les mêmes résultats ont été observés pour le mutant rTrfbL2-D77A et les deux mutants ont donc été considérés instables pour la construction dans le vecteur pET32-TEV. Afin de vérifier que l'instabilité n'est pas due à la nature de l'étiquette ajoutée en N-ter de la protéine, ce mutant a été produit dans les vecteurs pET-TEV et pProEx Htb qui ajoutent une étiquette plus petite uniquement constituée d'une séquence poly-histidine. Les mêmes précipitations ont été observées. Cela laisse supposer que les précipités observés dans les

extraits solubles des mutants produits sans étiquette dans le vecteur pET29a sont aussi dus à rTrfbL2 mutée. Il a donc été conclu que la mutation D77A déstabilisait la protéine. L'analyse structurale montre qu'un des oxygènes de l'aspartate 77 interagit avec la fonction OH de la tyrosine 95 et que le deuxième oxygène interagit avec une molécule d'eau (**Figure IV-10**). L'aspartate et la tyrosine sont conservées chez les protéines homologues et ces interactions sont retrouvées dans l'ensemble des structures et pourraient donc avoir un fort impact sur la stabilité. De plus, cette interaction fait le lien entre une des boucles du motif hélice-tour-hélice et le brin G et pourrait donc être nécessaire au bon repliement de TrfbL2. Enfin, en absence de GlcNAc, l'aspartate 77 interagirait avec des molécules d'eau à la place du GlcNAc comme c'est le cas dans les structures apo de ABL, BL, SRL et XCL car il s'agit d'une zone exposée au solvant. Ces interactions sont supprimées dans le cas d'une mutation en alanine.

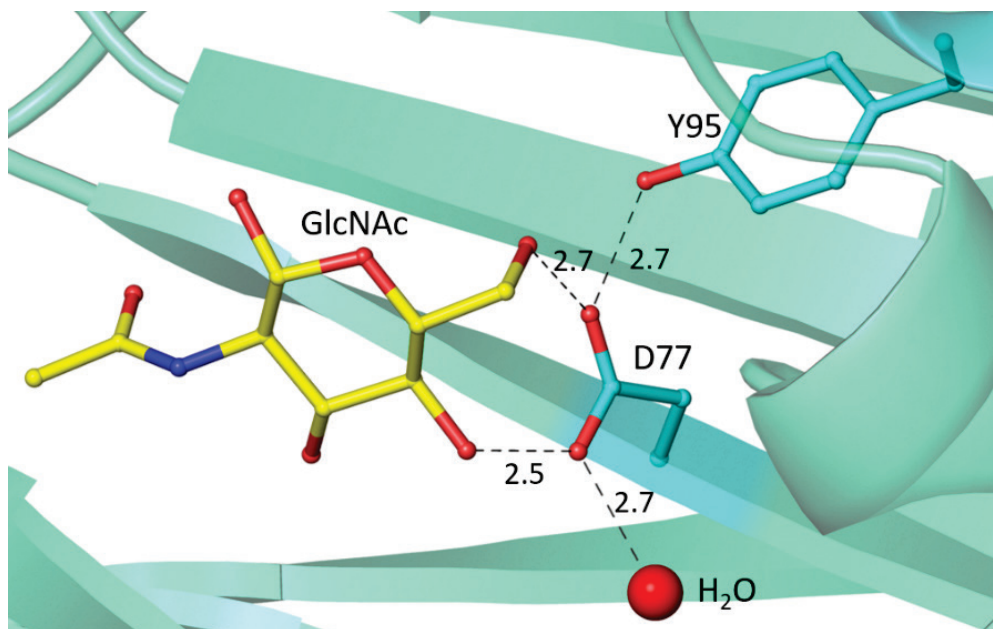


Figure IV-10. Description des interactions de la chaîne latérale de l'aspartate 77 de rTrfbL2.

Les distances d'interaction sont indiquées en Å.

1.2.4 Caractérisation du mutant R100G : stabilité et activité

Compte tenu des problèmes de stabilité rencontrés avec la mutation D77A, le troisième mutant R110G a ensuite été étudié, toujours pour diminuer l'affinité du site 2 d'interaction pour le GlcNAc. Ce mutant a été préalablement cloné dans le vecteur pET29a puis produit dans des cellules Rosetta pLysS selon le même protocole que pour rTrfbL2 sauvage. Afin d'évaluer rapidement l'intérêt de cette mutation pour empêcher la fixation du GlcNAc sur le deuxième site, la méthode de purification sur colonne d'affinité au GlcNAc a été choisie. L'analyse de la purification sur Gel SDS-PAGE est présentée **Figure IV-11**. Une bande

apparaît sous 15 kDa et semble correspondre à rTrfbL2-R100G puisqu'elle se trouve au même niveau que la bande de migration de rTrfbL2-WT. On remarque que plus de la moitié de rTrfbL2-R100G se trouve dans la fraction insoluble, mais cela est aussi le cas lors de la production de rTrfbL2 sauvage. Pour ce qui est de l'extrait soluble, la majeure partie de rTrfbL2-R100G ne se fixe pas sur la colonne d'affinité au GlcNAc et est éluée lors du lavage réalisé avec 20 mM Tris pH 7,5 + 100 mM NaCl. Cela semble indiquer que cette mutation permet de diminuer l'affinité de rTrfbL2 pour le GlcNAc comparé au mutant Y111A. De plus, une faible partie de rTrfbL2-R100G est retrouvée dans l'élué, dont les dernières fractions 5, 6 et 7 paraissent pures. Cependant à l'issue de la dialyse, une partie de rTrfbL2-R100G a précipité, suggérant une moins bonne stabilité de rTrfbL2-R100G comparé à la protéine sauvage. (**Figure IV-11.B**). Les fractions 5, 6 et 7 ont donc été concentrées puis dialysées afin d'analyser la stabilité de ce mutant.

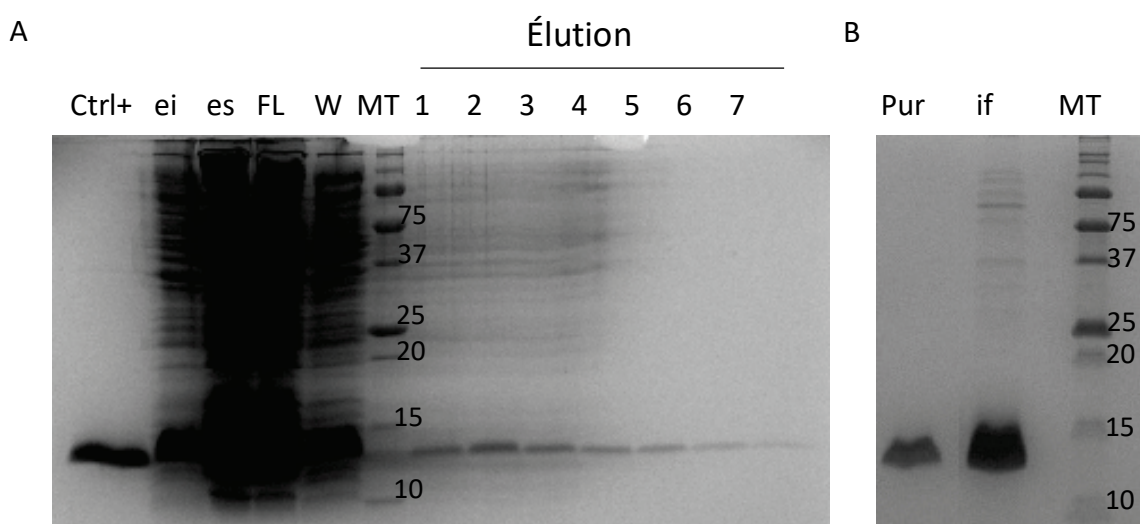


Figure IV-11. A. Gel SDS-PAGE tris-glycine 15% de la purification de rTrfbL2-R100G sur colonne d'affinité au GlcNAc.

Ctrl+ = rTrfbL2-WT pure ; ei = extrait insoluble ; es = extrait soluble ; FL = fraction non retenue ; W = lavage ; MT = marqueur de taille en kDa. B. Pur = fractions 5, 6 et 7 concentrées ; if = précipité de rTrfbL2-R100G.

La stabilité de rTrfbL2-R100G a été analysée par un test TSA en même temps que rTrfbL2-WT et rTrfbL2-Y111A pour comparaison (**Figure IV-12**). Ce test confirme que rTrfbL2-WT et Y111A ont une stabilité thermique similaire avec une température de fusion proche de 80°C et une dénaturation en une seule étape. Ce n'est pas le cas de rTrfbL2-R100G qui se dénature en deux étapes avec un premier T_m de 60°C et un deuxième T_m de 79°C. Cela semble indiquer que la mutation R100G fragilise la stabilité de rTrfbL2. La dénaturation en deux étapes peut s'expliquer par la tétramérisation de rTrfbL2 avec une première

séparation des dimères (ou monomères) suivie de la dénaturation de chaque monomère par exemple.

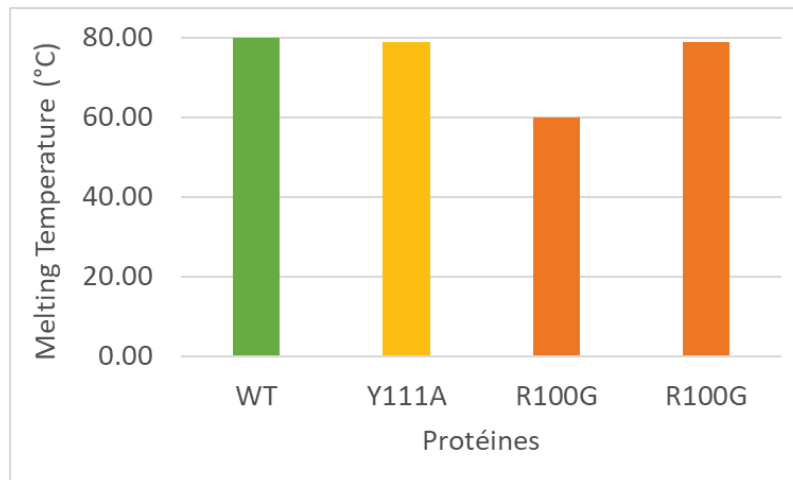


Figure IV-12. Test de stabilité avec rTrfbL2-WT, rTrfbL2-Y111A et rTrfbL2-R100G.

Les tampons ont été dilués à 0.1M. rTrfbL2 a été utilisée à 0.1 mg.mL^{-1} dans du sypro orange 5x. Les mesures ont été réalisées sur un appareil de QRT-PCR (Mini-Opticon Biorad).

L'activité du mutant rTrfbL2-R100G a ensuite été vérifiée par des tests d'hémagglutination de globules rouges de lapin (**Figure IV-13**). Ce test permet de vérifier rapidement que la protéine est capable de reconnaître ses ligands en surface des globules rouges. Les résultats montrent que rTrfbL2-R100G a la même capacité d'hémagglutination que la protéine sauvage avec une concentration limite d'hémagglutination autour de $0.5 \text{ }\mu\text{g.mL}^{-1}$, ce qui indique que ce mutant est actif et qu'il est donc probablement bien replié.

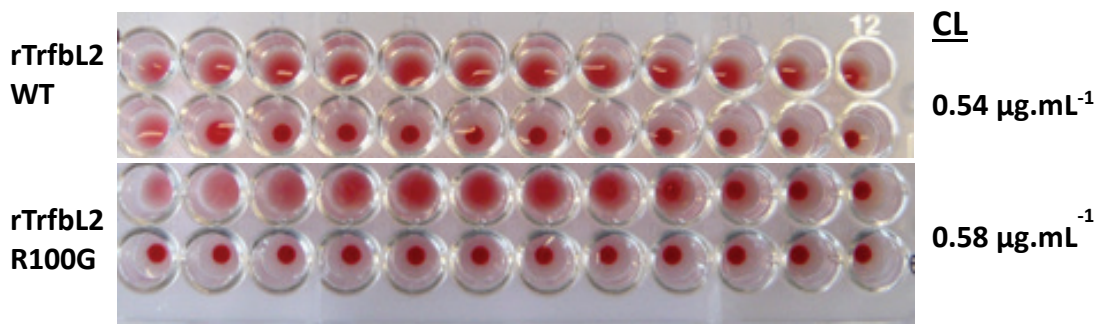


Figure IV-13. Comparaison des capacités d'hémagglutination de rTrfbL2-WT et de rTrfbL2-R100G.

CL = concentration limite de lectine pour laquelle il y a encore hémagglutination. rTrfbL2-WT à 2.2 mg.mL^{-1} et rTrfbL2-R100G à 0.15 mg.mL^{-1} dans le premier puits.

Malheureusement, malgré deux T_m supérieures à 50°C et une bonne capacité d'hémagglutination, le mutant rTrfbL2-R100G n'est pas stable dans le temps et précipite après quelques jours à 4°C . S'ajoute à cela le très faible rendement de purification sur colonne

d'affinité au GlcNAc (environ 150 µg de protéine pure par litre de culture). Des tentatives de purification sur colonne échangeuse d'ions ont été réalisées pour essayer d'augmenter le rendement, mais comme pour les mutants rTrfbL2-D77A et rTrfbL2-Y111QD77A, cette méthode ne permet pas de purifier rTrfbL2-R100G.

L'instabilité de ce mutant peut aussi trouver une explication dans l'analyse de la structure cristallographique. En effet, le pont salin entre le noyau guanidinium de l'arginine 100 et le glutamate 94 joue probablement un rôle dans la stabilité de la structure et ne peut plus exister après mutation de l'arginine en glycine (**Figure IV-14**). Il en est de même pour les interactions hydrogènes entre les fonctions NH1 et NH de l'arginine et l'oxygène du glutamate. L'analyse des séquences homologues montre la conservation de l'arginine 100 et la présence d'un glutamate ou d'une glutamine en position 94 créant l'interaction hydrogène avec la chaîne latérale de l'arginine 100. L'arginine 100 forme une partie de la boucle reliant les deux hélices α tandis que le glutamate est impliqué dans la formation d'une des deux hélices. Les interactions existantes entre ces deux acides aminés sont donc probablement impliquées dans la stabilité du motif hélice-tour-hélice de TrfbL2. De plus, la glycine est caractérisée par son absence de carbone asymétrique et la mutation pourrait donc apporter de la flexibilité supplémentaire. L'ensemble de ces modifications peuvent avoir un impact non-négligeable dans la stabilité de rTrfbL2.

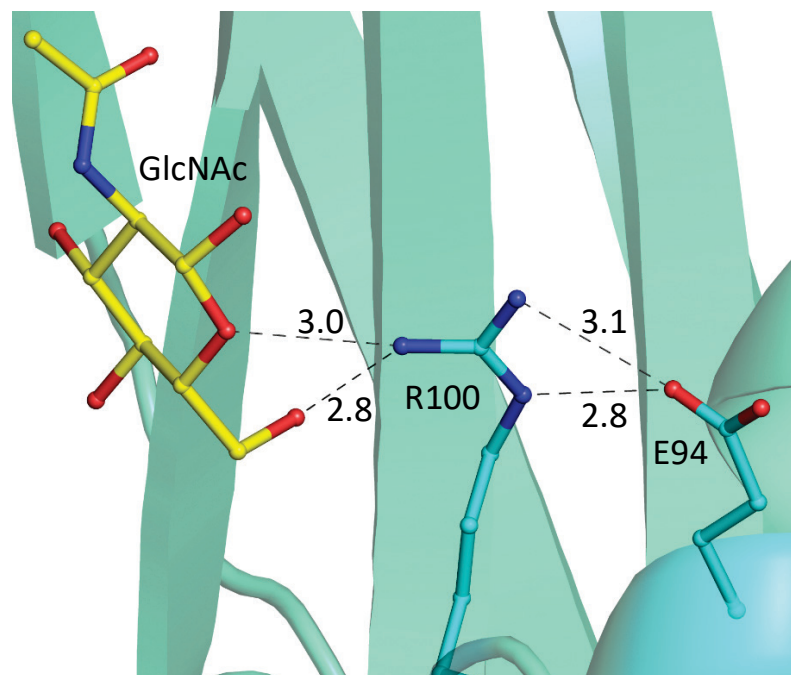


Figure IV-14. Description des interactions de la chaîne latérale de l'arginine 100 de rTrfbL2.

Les distances d'interaction sont indiquées en Å.

Il serait intéressant de pouvoir réaliser des études de modélisation moléculaire afin de définir si l'on peut trouver d'autres mutations pour l'aspartate 77 ou l'arginine 100 qui soient moins déstabilisantes, tout en empêchant la fixation de GlcNAc.

Le mutant rTrfbL2-Y111A pouvant être obtenu facilement avec une affinité très faible pour le GlcNAc, les études biologiques sur cellules et tissus cancéreux ont été réalisées avec ce mutant afin d'évaluer si la reconnaissance du GlcNAc peut être considérée comme négligeable et évaluer son potentiel pour détecter l'antigène TF en cancérologie (*voir Chapitre IV. 4*).

Chapitre 2.

**Reconnaissance de la core-
fucosylation : rPhoSL**

2. Reconnaissance de la core-fucosylation : rPhoSL

PhoSL a été précédemment étudiée sous sa forme native, et purifiée à partir d'extraits du champignon *Pholiota squarrosa* (Kobayashi et al., 2012). Cette étude a démontré une reconnaissance stricte de la core-fucosylation par PhoSL, d'où un fort intérêt en tant qu'outil de marquage. L'utilisation de la protéine native n'étant pas adaptée, l'objectif de ce travail a été de développer la forme recombinante de rPhoSL afin d'assurer une bonne reproductibilité, une qualité et quantité définies, de faciliter sa production sans être dépendant de l'approvisionnement en champignons, et de caractériser le mécanisme de reconnaissance de cette protéine ainsi que de résoudre sa structure 3D.

2.1 Production et purification

Étant donné qu'un pont disulfure est prédit pour PhoSL, nous avons envisagé deux possibilités pour sa production sous forme recombinante (rPhoSL) : soit réaliser la production de la lectine dans le périplasme des cellules qui est un environnement oxydant, soit utiliser une souche modifiée d'*E. coli* pour laquelle la voie des réductases a été neutralisée dans le cytoplasme. Les deux méthodes ont guidé le choix du plasmide et la nature de la souche à utiliser puisqu'ils peuvent avoir un impact sur le bon repliement des protéines ou encore les rendements de production. Des résultats préliminaires ont montré que les taux de production étaient équivalents dans les deux cas pour rPhoSL. La production dans le périplasme a par la suite été retenue car elle permet de garantir la formation des ponts disulfures et de faciliter la qualité de la purification en raison d'un taux plus faible de protéines d'*E. coli* dans le périplasme.

Le gène de rPhoSL a donc été cloné dans le vecteur pET39-TEV qui permet d'ajouter une étiquette contenant l'enzyme DsbA présentant un signal de localisation vers le périplasme et une séquence polyhistidine clivable avec la protéase TEV en N-ter de la protéine (étiquette d'environ 29 kDa, voir **Chapitre III**). L'expression est réalisée dans la souche BL21 Star (DE3) d'*E. coli* suivie d'une extraction périplasmique des protéines de façon à ne pas être contaminé par les protéines cytoplasmiques d'*E. coli*. rPhoSL a été purifiée par chromatographie d'affinité sur une résine IMAC grâce à la présence de la séquence polyhistidine. Les résultats ont été analysés sur gel SDS-PAGE Tris-glycine 15 %. Les bandes correspondant aux tailles attendues pour l'ensemble protéine-étiquette ont pu être observées vers 35 kDa (**Figure IV-15**). Après élimination de l'imidazole par dessalage, l'étiquette est coupée sur la nuit à 19°C par la protéase TEV produite au laboratoire. La protéine est ensuite

redéposée sur colonne IMAC et la fraction non retenue contenant la protéine sans séquence poly-histidine est récupérée. Afin d'améliorer la pureté, rPhoSL a subi une troisième étape de purification sur une colonne échangeuse d'anions DEAE. En effet, rPhoSL a un point isoélectrique (théorique) inférieur à 5 et est donc chargée négativement dans un tampon à pH 8 comme celui utilisé. Les fractions sont finalement analysées sur un gel SDS-PAGE Tris-tricine 16.5 % qui permet de voir les protéines de poids moléculaire inférieur à 10 kDa (**Figure IV-15**). Après toutes ces étapes, la protéine est pure à environ 95 %, permettant d'envisager sa caractérisation biochimique et structurale. Elle a été concentrée dans 25 mM Mes pH 6 et conservée à 4°C. Le rendement de production et de purification est d'environ 1 mg par litre de culture.

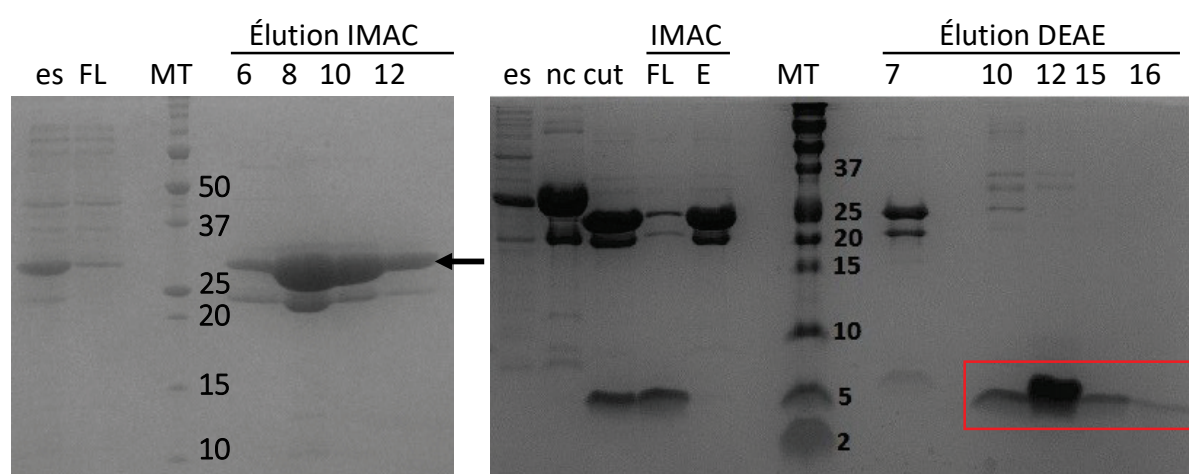


Figure IV-15. Gels SDS-PAGE des deux étapes de purification de rPhoSL sur IMAC et DEAE.

Gel SDS-PAGE Tris glycine 15% (gauche) ou Tris-tricine 16.5 % (droite). es = extrait soluble ; FL = fraction non retenue ; MT = marqueur de taille en kDa ; E : élution sur colonne IMAC. La flèche indique la taille attendue pour rPhoSL-étiquette.

Il est important de souligner qu'après clonage dans le vecteur pET39-TEV puis coupure de l'étiquette par la protéase TEV, trois acides aminés « GAM » ont été ajoutés en N-terminal de rPhoSL alors qu'ils n'existent pas dans la séquence d'origine (**Figure IV-16**).

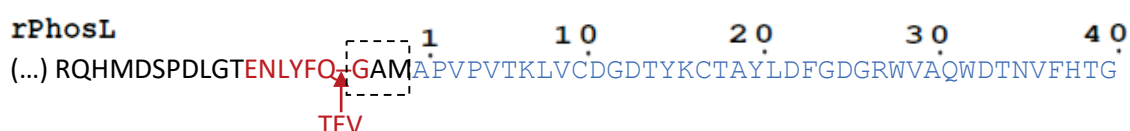


Figure IV-16. Ajout des trois acides aminés « GAM » lors des étapes de clonage dans le vecteur Pet39-TEV.

Description de la coupure de l'étiquette en N-terminal : séquence de l'étiquette en noir, séquence reconnue par la TEV en rouge avec son site de coupure désigné par une flèche. Séquence de PhoSL en bleu. Les 3 acides aminés GAM ajoutés après la coupure à la TEV sont entourés en pointillés.

2.2 Analyse de l'état d'oligomérisation, de la stabilité, et de l'activité

Une fois que la lectine rPhoSL a été purifiée en quantité suffisante, elle a été analysée pour vérifier son état d'oligomérisation, son activité et sa stabilité.

Des expériences de filtration sur gel ont montré que rPhoSL formait des trimères en solution. Son activité a ensuite été vérifiée par des tests d'hémagglutination de globules rouges de lapin (**Figure IV-17**). rPhoSL est capable d'hémagglutiner les globules rouges, ce qui indique à la fois qu'elle est active et bien repliée, mais aussi qu'elle est multimérique. Le test a été réalisé avec ou sans traitement préalable des globules rouges avec de la trypsine, une protéase. rPhoSL montre une capacité d'hémagglutination environ 3 fois supérieure en présence de trypsine. Celle-ci facilite donc vraisemblablement l'accès de la lectine à certains ligands peu exposés.

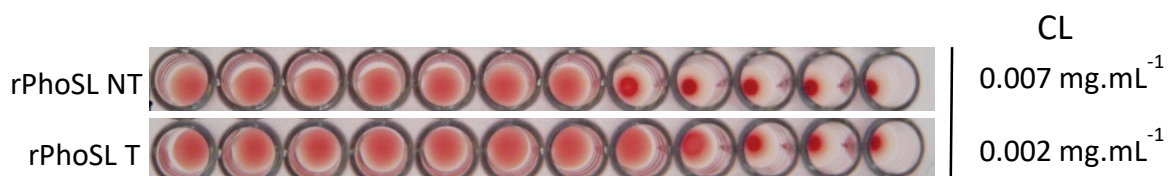


Figure IV-17. Test d'hémagglutination de rPhoSL sur globules rouges de lapin avec du sang préalablement traité (T) ou non (NT) à la trypsine.

CL = concentration limite de lectine pour laquelle il y a encore hémagglutination. rPhoSL à 0.45 mg.mL⁻¹ dans le premier puits.

Enfin, la stabilité de rPhoSL a été étudiée au cours de tests de TSA (Thermal Shift Assay). Les tests réalisés avec rPhoSL sur 26 tampons de pH 5 à pH 10 ont montré que cette lectine était stable dans tous les tampons testés avec des températures de fusion supérieures à 60 degrés (**Figure IV-18**). Sa température de fusion maximale, de 68 °C, est atteinte pour des tampons à pH acide. Cette expérience a permis de définir le MES pH 6 comme tampon de conservation pour cette lectine, du fait de son point isoélectrique théorique égal à 4.66. Afin de déterminer si la concentration en NaCl a un impact sur la stabilité de la lectine, des tests de stabilité avec ces mêmes tampons et 0 à 100 mM de NaCl ont aussi été réalisés. Ils n'ont montré aucun effet du NaCl sur la stabilité de rPhoSL.

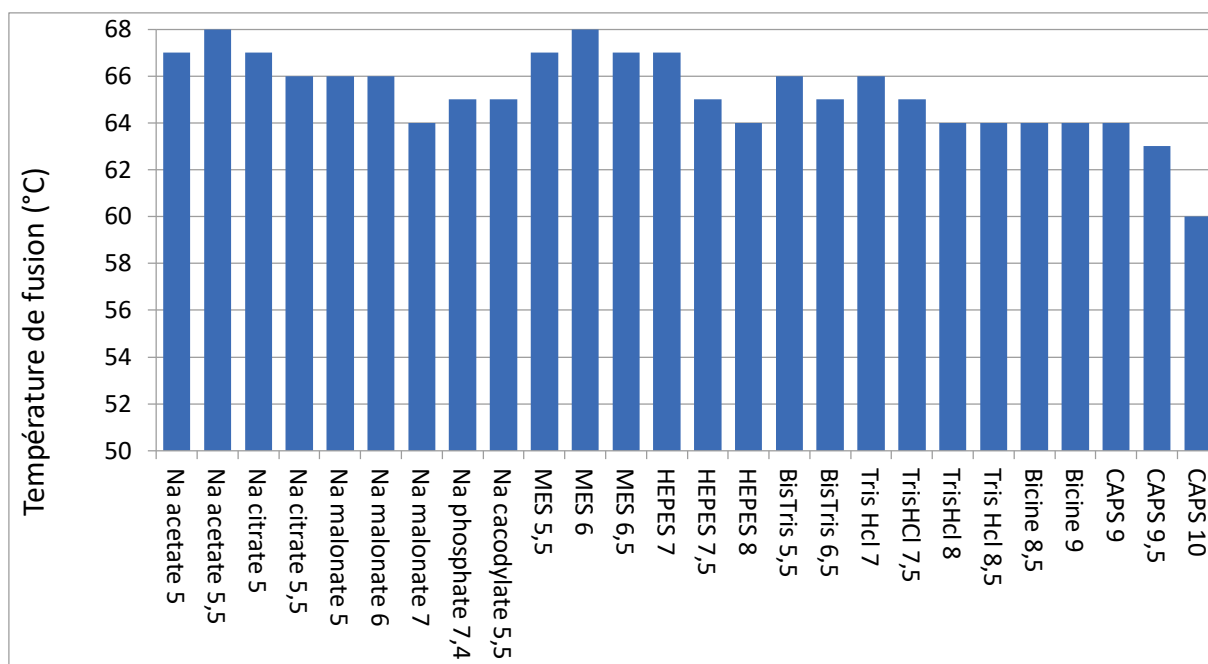


Figure IV-18. Test TSA avec rPhoSL et 26 tampons de pH 5 à PH 10.

Les tampons ont été dilués à 0.1M. rPhoSL a été utilisée à 0.1 mg.mL⁻¹ dans du Sypro orange 5x. Les mesures ont été réalisées sur un appareil de QRT-PCR (Mini-Opticon Biorad).

2.3 Étude de l'interaction lectine-sucre

2.3.1 Analyse de la spécificité sur puces à sucres (glycan array)

La protéine native avait précédemment été décrite comme étant strictement spécifique de la core-fucosylation des N-glycanes. Dans le but de vérifier que la protéine sous forme recombinante présente cette même propriété essentielle pour son utilisation en tant qu'outil de marquage, sa spécificité a été analysée sur une puce constituée de 600 sucres présents à la surface des cellules de mammifères (Consortium for Functional Glycomics, CFG).

Les résultats confirment que rPhoSL reconnaît uniquement les N-glycanes core-fucosylés en $\alpha(1,6)$ (**Figure IV-19**). Dès lors que la core-fucosylation est présente, rPhoSL reconnaît tous les types de N-glycanes, bissectés ou non, quelque soit le nombre de branches et la nature des glycanes qui les composent.

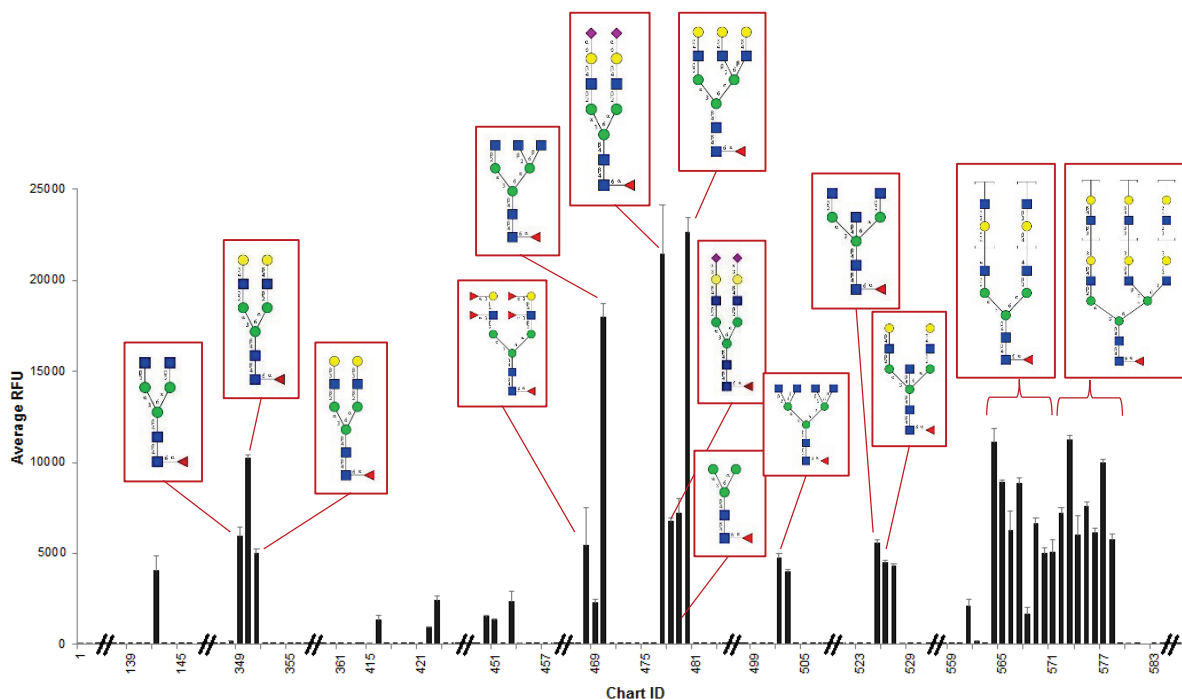


Figure IV-19. Résultats de l'étude de rPhoSL sur puce à sucres version 5.2.

rPhoSL ($5 \mu\text{g.mL}^{-1}$) a été marquée avec l'alexa 488. Seuls les sucres avec un signal RFU supérieur à 20 % du signal maximal ont été représentés.

2.3.2 Mesure de l'affinité par Titration Calorimétrique Isotherme (ITC)

L'objectif a ensuite été de déterminer l'affinité de rPhoSL pour des sucres core-fucosylés en $\alpha 1,6$ afin de caractériser la force des interactions entre la lectine et ses ligands. Ces premières mesures ont été réalisées en solution en utilisant la technique de titration calorimétrique isotherme qui a l'avantage d'être rapide à mettre en œuvre et qui ne nécessite pas de marquage préalable de la protéine.

Les mesures ont été réalisées avec du chitobiose core-fucosylé (N2F N-glycan, Carbosynth) et le N-glycane 1 (octasaccharide). Les valeurs de K_D obtenues sont de $180 \mu\text{M}$ pour le chitobiose core-fucosylé et de $105 \mu\text{M}$ pour le N-glycane 1 (**Figure IV-20**). Le fait qu'on ait une affinité environ 1,8 fois meilleure avec le N-glycane 1 peut être due à des interactions supplémentaires avec les monomères absents du chitobiose core-fucosylé.

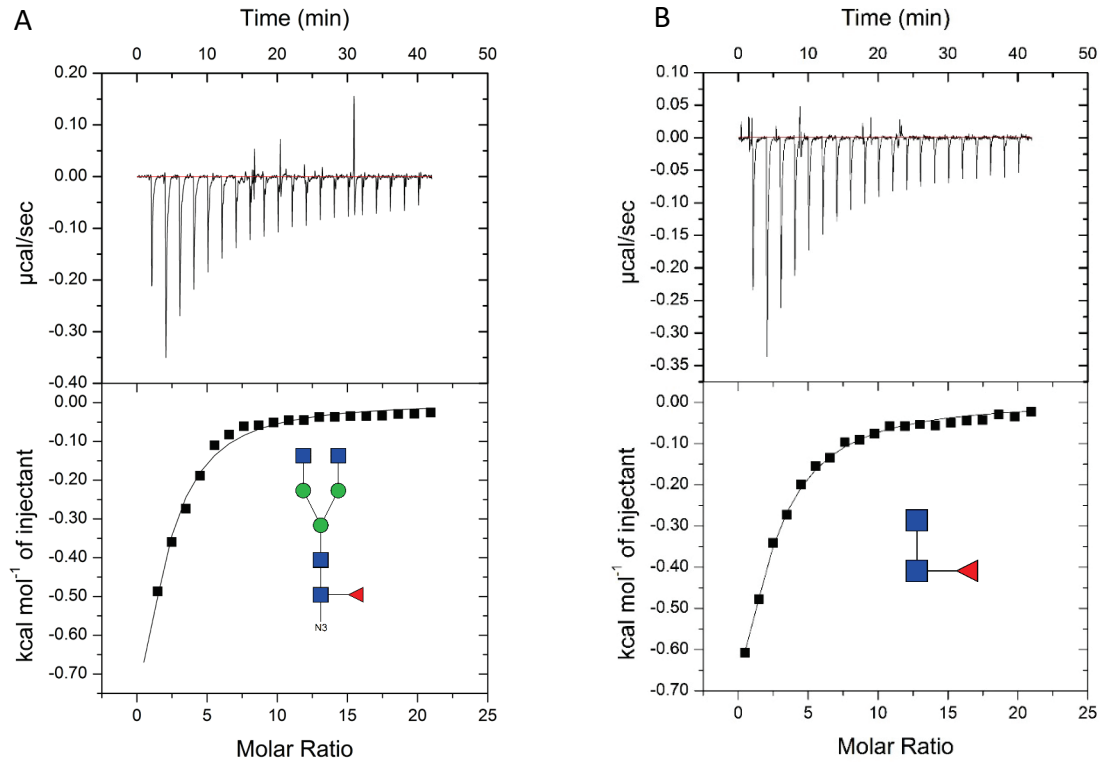


Figure IV-20. Résultat des mesures d'affinité de rPhoSL pour le N-glycane 1 (A) et le chitobiose core-fucosylé (B) par ITC.

rPhoSL à 50 μM a été titrée avec 10 mM de N-glycane 1 (code GlyTouCan G80858MF) et 5 mM de chitobiose core-fucosylé (code GlyTouCan G21290RB) à 25°C. La protéine et les sucres ont été préparés dans 25 mM MES pH 6.

Ces valeurs d'affinité sont cependant 100 fois plus faibles que celles observées pour la protéine extraite du champignon ($3.3 \cdot 10^{-6}$ M, mesuré par chromatographie d'affinité frontale ; (Kobayashi et al., 2012)) et pour SL2-1, une lectine homologue ($1.59 \cdot 10^{-6}$ M, mesuré par ITC ; (Vainauskas et al., 2016)). Ces différences d'affinité peuvent être liées au fait que les techniques utilisées ne sont pas les mêmes (entre chromatographie d'affinité frontale et ITC) ou bien aux différences existantes au sein des séquences protéiques : acides aminés supplémentaires en N-ter de rPhoSL comparé à PhoSL native, et différences avec la séquence de SL2-1 (*Figure IV-21*).

	1	10	20	30	40							
rPhoSL	GAM	APVPVT	KL	LV	CDGDTY	KCTAY	LD	FGDGRWVAQW	DT	NVFHT	G
PhoSL	...	APVPVT	KL	LV	CDGDTY	KCTAY	LD	FGDGRWVAQW	DT	NVFHT	G
SL2-1a	...	APVPVT	SL	LN	CDGDTY	TCTA	GL	RF	FGDGRWVAQW	NN	NVFHQ	SPASQQRA...ARSADAPRGAA
SL2-1b	...	APVPVT	TL	LT	CDGDTY	KCTA	GL	QF	FGDGRWVAQW	NN	NVFHQ	AMANAVPTASAKTVKAPREAA
SL2-1c	...	APVPVT	AL	LN	CDGDTY	KCTA	GL	LF	FGDGRWVAQW	DA	SVFHQ	S.....

Figure IV-21. Alignement de la séquence de rPhoSL avec celles de PhoSL native et des trois répétitions de SL2-1.

SL2-1 est constituée de trois répétitions homologues à la séquence de rPhoSL dénommées ici SL2-1a, SL2-1b et SL2-1c.

2.3.3 Tests en SPR (Résonance Plasmonique de Surface) et BLI (Bio-Layer Interferometry)

Les mesures d'affinité par ITC étant effectuées en solution, il était ensuite intéressant de mesurer l'affinité de rPhoSL pour des surfaces core-fucosylées. En effet, les lectines sont souvent multivalentes, et cela se traduit généralement par une affinité fonctionnelle plus élevée que leur affinité intrinsèque lorsque la mesure est réalisée sur une surface recouverte de ligands, comme c'est le cas en SPR. L'affinité fonctionnelle, aussi appelée avidité, correspond à la force globale de l'interaction, contrairement à l'affinité intrinsèque qui correspond à la force d'une seule interaction monovalente (Rheinnecker et al., 1996). Pour une lectine bivalente par exemple, une fois que l'interaction sur le premier site a eu lieu, la liaison sur le second site va être facilitée par la proximité entre la lectine et les ligands. De plus, la vitesse de dissociation est ralentie car la lectine a tendance à se re-fixer à la surface (Shinohara et al., 1997). De ce fait, l'affinité mesurée sur une surface peut être jusqu'à 10^4 fois plus élevée qu'en solution (Monsigny et al., 2000), et les faibles affinités intrinsèques des lectines sont alors compensées par leur multivalence.

La mesure de l'interaction de rPhoSL sur une surface permettrait donc de mimer davantage les interactions biologiques sur cellules et tissus. Les deux techniques de SPR (Surface Plasmon Resonance) et de BLI (Bio Layer Interferometry) ont donc été testées dans ce but. Pour cela, de la thyroglobuline biotinylée qui possède des sucres core-fucosylés, ligands de rPhoSL (Ronin et al., 1986), a été fixée sur le capteur (BLI) ou la puce (SPR) via une liaison streptavidine-biotine. rPhoSL a ensuite été ajoutée à différentes concentrations pour mesurer l'interaction. Les premières expériences ont malheureusement révélé que bien que rPhoSL semblait se fixer à la thyroglobuline, elle se fixait également sur le capteur contrôle sans thyroglobuline. Ces interactions non-spécifiques sont telles qu'après soustraction du signal contrôle, le signal résultant est soit négatif, soit nul. Aucune détermination d'affinité n'a donc été possible puisqu'il est probable que ces interactions non-spécifiques existent aussi sur le capteur de mesure où est fixée la thyroglobuline. Plusieurs tests ont alors été mis en place pour éviter ces interactions non-spécifiques tel que l'ajout sur le capteur de molécules avec lesquelles rPhoSL ne devrait théoriquement pas interagir. C'est le cas des protéines LecA et LecB biotinylée ou du biot-PAA-GlcNAc. Cela a conduit dans certains cas à une faible baisse des interactions non-spécifiques sur le capteur contrôle, mais cela n'était pas d'ampleur suffisante pour espérer mesurer une affinité de façon précise. Suspectant des interactions non-spécifiques avec la streptavidine fixée sur les capteurs et les puces, la thyroglobuline a ensuite

été fixée directement par ses amines primaires sur des capteurs ou puces sans streptavidine. Mais cela n'a pas donné de meilleurs résultats. Des discussions avec des experts nous ont confirmé que les interactions non spécifiques étaient un problème récurrent pour ce type d'expérience, et qu'il était parfois difficile d'y palier, comme ici pour rPhoSL. Ces tests n'ont donc pas été poursuivis au cours de ce travail.

2.3.4 Tests d'interaction avec des glycoprotéines en ELLA (Enzyme-linked lectin assay)

Toujours dans le but de mieux caractériser les interactions entre rPhoSL et ses ligands, des tests ont ensuite été réalisés sur des glycoprotéines afin d'évaluer la capacité de la lectine à reconnaître les ligands core-fucosylés dans un environnement plus complexe. La technique ELLA a été utilisée car elle permet de visualiser très facilement l'interaction entre des lectines et des glycoprotéines fixées sur une plaque 96 puits. Des glycoprotéines dont la glycosylation est connue ont été testées : la fétuine, la thyroglobuline et les IgG- κ qui sont core-fucosylées ((Aich et al., 2013; Ronin et al., 1986; Berkel et al., 2009), ainsi que l'ovalbumine et la conalbumine qui ne le sont pas (Schneider and Nägele, 2014; Tai et al., 1975). Les glycoprotéines ont été déposées au fond des puits et la lectine préalablement biotinylée a été ajoutée après des étapes de lavage et de saturation. La révélation a été faite par coloration avec de la streptavidine-peroxidase associée à de l'OPD. rPhoSL reconnaît uniquement les glycoprotéines possédant la core-fucosylation (fétuine, thyroglobuline, IgG- κ) et pas l'ovalbumine et la conalbumine, **Figure IV-22**. Ces résultats prouvent que rPhoSL est capable d'interagir avec ses ligands au sein des glycoprotéines, avec la même spécificité qu'en solution. Cela est encourageant dans l'objectif d'utiliser cette lectine pour la reconnaissance de glycoprotéines core-fucosylées de sérums de patients par exemple.

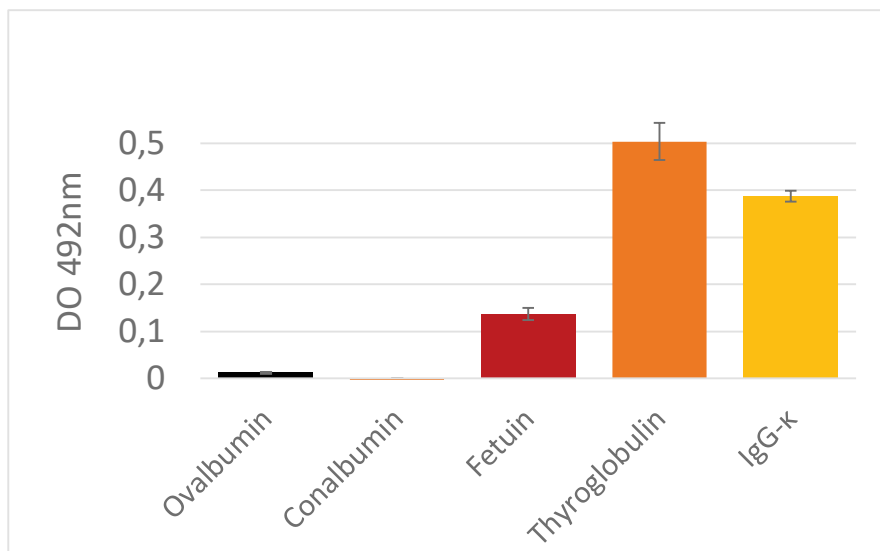


Figure IV-22. Fixation de rPhoSL biotinylée su des glycoprotéines par ELLA. L’ovalbumine et la conalbumine ne sont pas core-fucosylées. La fétuine, la thyroglobuline et l’IgG-κ sont core-fucosylées.

2.4 Étude des propriétés anti-adhésives de rPhoSL sur des micro-organismes

De nombreux micro-organismes, pathogènes ou non, ont la propriété d’adhérer à des surfaces, ce qui leur permet ensuite de les coloniser. Cela est à l’origine de problèmes de santé humaine comme les infections intestinales dues à *Salmonella*, la listériose due à *Listeria monocytogene* ou encore des maladies nosocomiales comme celles dues à *Staphylococcus aureus*. Certaines lectines, de par la reconnaissance d’oligosaccharides en surface de ces micro-organismes, vont avoir un effet barrière et empêcher l’adhésion de ceux-ci sur des surfaces.

Bactéries	Champignons	Levures
<i>Escherichia coli</i>	<i>Aspergillus flavus</i>	<i>Candida intermedia</i>
<i>Bacillus cereus</i>	<i>Aspergillus ochraceus</i>	
<i>Bacillus subtilis</i>	<i>Pichia anomala</i>	
<i>Campylobacter jejuni</i>		
<i>Salmonella</i>		
<i>Pseudomonas fragi</i>		
<i>Listeria monocytogenes</i>		

Tableau IV-2. Tableau des différents micro-organismes utilisés pour l’étude des propriétés anti-adhésives de rPhoSL.

Dans le but de tester la capacité de rPhoSL à empêcher l'adhésion de micro-organismes, j'ai réalisé un séjour de deux semaines à Ljubljana (Slovénie) dans le cadre d'un échange bilatéral financé par le programme Protéus avec Campus France. Des tests ont été réalisés sur onze micro-organismes (*Tableau IV-2*). Le principe a été de mettre les micro-organismes en présence, ou non, de rPhoSL avant de les incuber sur des plaques 96 puits. Après lavage, seuls les micro-organismes qui étaient toujours capables d'adhérer sur la plaque, malgré la présence de rPhoSL, restaient fixés. Les colonies ont ensuite été comptées après transfert en boîte de pétri pour rendre compte de la quantité de micro-organismes dont l'adhérence a été impactée par la lectine. Les détails du protocole sont décrits *chapitre III. 7*. Aucune diminution significative du nombre de colonies n'a été observée lors de la pré-incubation des micro-organismes avec rPhoSL, ce qui indique que cette protéine n'a pas d'activité anti-adhésive sur les organismes testés.

2.5 Étude structurale

Au début de ce travail, il n'existait aucune structure protéique permettant de décrire ni comprendre la reconnaissance spécifique de la core-fucosylation. Ces informations sont essentielles pour révéler les déterminants protéiques et glucidiques qui gouvernent cette spécificité et ainsi pouvoir optimiser la reconnaissance, ou trouver de nouveaux candidats par homologie. Une étude structurale a donc été menée dans le but de résoudre la structure de rPhoSL en absence et en présence de ligands.

2.5.1 Cristallogénèse

Afin d'obtenir des cristaux de rPhoSL, un criblage de conditions de cristallogénèses a été réalisé au laboratoire ou à l'aide d'un robot de cristallisation (HTX, EMBL, Grenoble). La lectine a été utilisée à des concentrations comprises entre 5.5 et 6.3 mg.mL⁻¹ dans un tampon

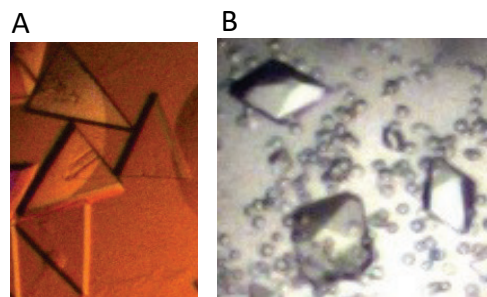


Figure IV-23. Photos des cristaux de rPhoSL (A) et de rPhoSL-ligands (B).

25 mM MES PH6 et pré-incubée ou non avec 1 mM de N-glycane fucosylé en α 1,6. Des cristaux de forme triangulaire ont été obtenus pour rPhoSL native avec 18 % butane-1,4-diol, 300 mM acétate de zinc et 0,1 M Imidazole-HCl pH 6-7. Des cristaux en forme de diamants ont été obtenus en complexe avec différents N-glycanes avec 3 à 3.4 M de malonate de sodium pH 5 (*Figure IV-23*).

2.5.2 Résolution de la structure de rPhoSL en absence de ligand

Les données de rPhoSL en absence de ligand ont été collectées sur la ligne BM14 à l'ESRF (European Synchrotron Radiation Facility) à Grenoble en utilisant un détecteur MARCCD. Les clichés de diffraction ont été traités avec XDS (Kabsch, 2010) et tous les calculs suivants ont été réalisés avec la suite de logiciels CCP4i (Winn et al., 2011). Les statistiques de collecte et d'affinement sont résumées dans le *Tableau IV-3*.

	rPhoSL natif		
Données collectées			
Ligne de lumière	ESRF BM14		
Groupe d'espace	P2 ₁		
Dimensions maille a b c (Å)	28.2 67.7 31.0		
α β γ (°)	90 97.25 90		
Longueur d'onde	0.984		
Resolution (Å)	33.84-1.70 (1.73-1.70)		
Réflexions mesurées	50099		
Réflexions uniques	12666		
Rmerge	0.045 (0.349)		
Mean I / sd(I)	14.8 (3.0)		
Complétude (%)	99.5 (97.6)		
Multiplicité	4.0 (3.6)		
CC _{1/2}	0.997(0.892)		
Estimated mean FOM / Pseudo free CC (%)	0.639 / 68.2		
Affinement			
Résolution (Å)	33.84-1.70		
No. of work/ free Reflections	12732 / 665		
R _{work} / R _{free} (%)	17.1 / 21.1		
Rmsd bond lengths (Å)	0.0151		
Rmsd bond angles (°)	1.62		
Nb d'atomes	Chaîne A	Chaîne B	Chaîne C
Protéine	277	338	339
Eau	41	39	46
B-facteurs moyens	Chaîne A	Chaîne B	Chaîne C
Protéine	25	24.7	22.8
Ligand			
Autres hétéroatomes	36.9	39.5	34.2
Eau	36.8	37	36.8
Ramachandran plot (%)			
Allowed	100		
Favored	99.2		
Outliers	0		
code PDB	6EKE		

Tableau IV-3. Données de collecte et d'affinement pour la structure native de rPhoSL. Les valeurs entre parenthèses correspondent à la couronne de plus haute résolution. R_{free} a été calculé avec 5 % des réflexions.

La structure de rPhoSL a été résolue dans le groupe d'espace $P2_1$ à 1.7 Å par méthode SAD avec le logiciel ShelXC/D (Sheldrick, 2010) en utilisant le signal anomal du zinc (**Figure IV-24**). L'affinement des structures a été réalisé avec Refmac5.8 (Murshudov et al., 2011) et les corrections et constructions manuelles des modèles ont été réalisées dans COOT (Emsley et al., 2010). La qualité stéréochimique des modèles affinés a été validé sur le serveur wwPDB Validation : <http://wwpdb-validation.wwpdb.org>.

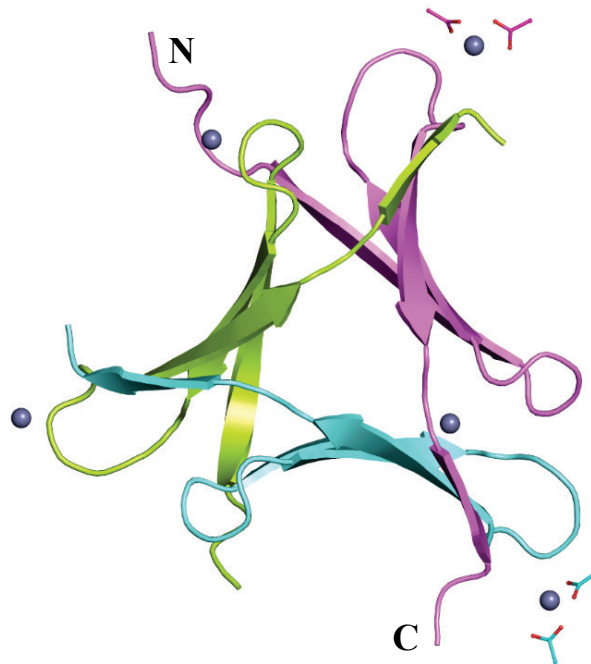


Figure IV-24. Représentation schématique de la structure rPhoSL native (β -prisme-III). Chaque couleur correspond à un monomère de rPhoSL. Les atomes de zinc sont représentés par des sphères grises et les ions acétate en visualisation de type bâtonnet.

rPhoSL est présente sous forme de trimères dans l'unité asymétrique et cinq atomes de zinc aident à la stabilisation des boucles de surface. Les monomères de rPhoSL forment des feuillets β antiparallèles à quatre brins légèrement incurvés qui ont été nommés brins A, B, C et D en allant du N au C-terminal. Les brins A et D sont à 90° l'un de l'autre (**Figure IV-25**). Les trois monomères sont arrangés entre eux selon un axe de symétrie trois pour former un β -prisme à 12 brins. Les brins β ne sont pas parallèles à l'axe du prisme mais subissent une rotation de 45° . De plus, le prisme est stabilisé par des interactions de type velcro suite à l'échange du brin D entre les différents monomères. D'autres contacts sont observés entre les extrémités des brins A, entre la cystéine 10 et la valine 3 ce qui contribue à la stabilisation des brins A (**Figure IV-25.C**). Le β -prisme présente une forme triangulaire équilatérale avec une torsion de 30° entre le bas (défini par le brin A de chaque feuillet) et le haut (formé par la boucle reliant les brins C et D). Son centre est majoritairement composé de résidus

hydrophobes tels que la valine 5 et la leucine 8 du brin A, le tryptophane 32 du brin C et le pont disulfure entre les cystéines 10 et 17 appartenant aux brins A et B (**Figure IV-25.D**).

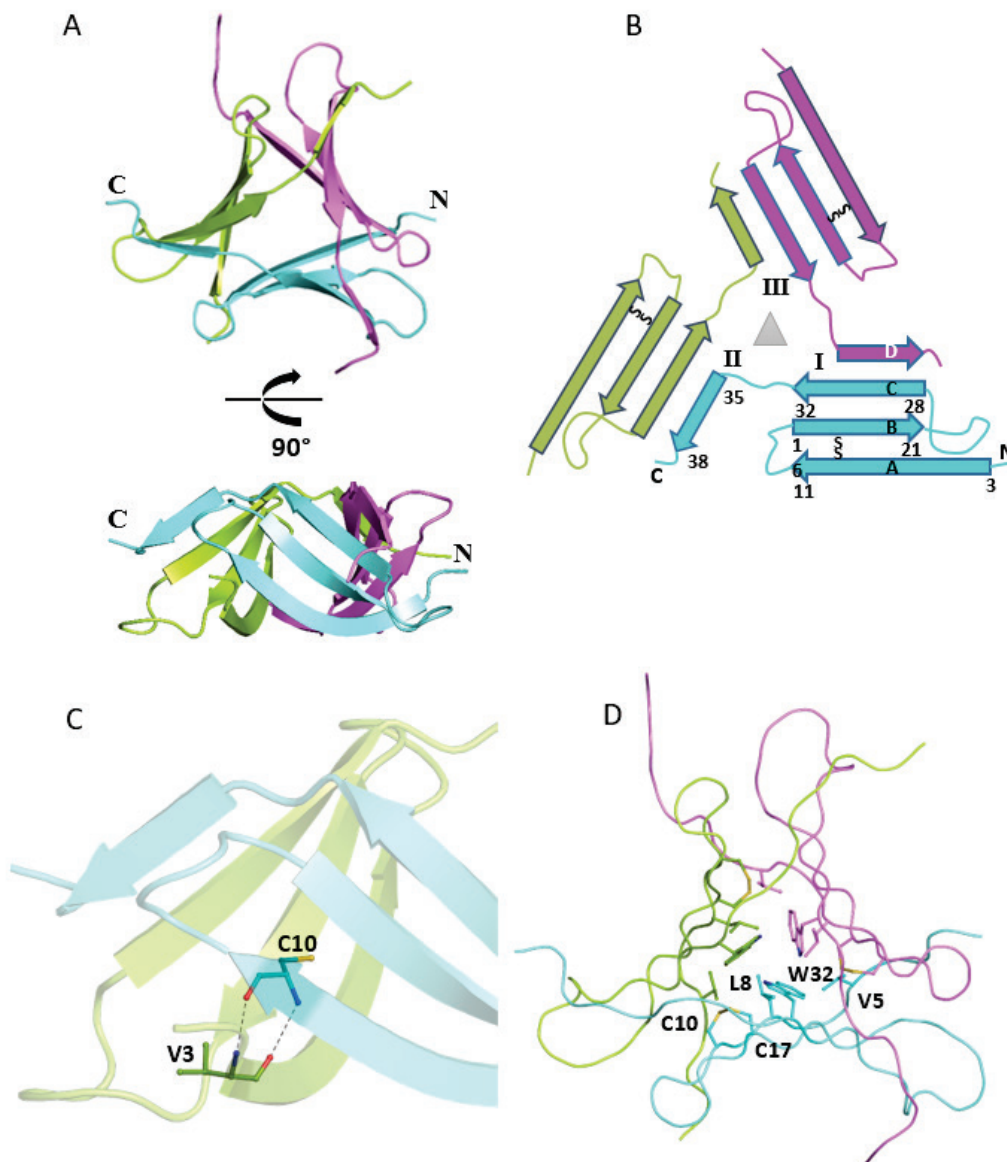


Figure IV-25. Description détaillée de la structure de rPhoSL.

A. Vues de dessus et de côté du β -prisme-III de rPhoSL. B. Diagramme topologique du trimère de rPhoSL. C. Zoom sur les interactions qui stabilisent le brin A entre deux monomères de rPhoSL. D. Résidus hydrophobes présents au centre du β -prisme III.

La conformation en 4 brins β adoptée par rPhoSL est un motif protéique commun que l'on trouve dans les motifs de type clé grecque, β -propeller, β -prisme II ou tonneaux- β (Fox et al., 2014). Le monomère de rPhoSL présente donc des homologies structurales avec les feuillettes de nombreuses protéines de type β -propeller (Holm and Rosenström, 2010). Mais pour le trimère complet, la recherche d'homologues structuraux avec PDBeFold (Krissinel and Henrick, 2004) n'a révélé que des homologies très éloignées avec le repliement de type β -

prisme II qui a été identifié pour la première fois pour l'agglutinine de *Galanthus nivalis* (GNA), spécifique du mannose (Hester et al., 1995). rPhoSL et GNA ne se superposent que pour un seul feuillet β car l'assemblage et l'agencement de leur brins β sont très distincts, et les feuillets β sont dirigés dans des directions opposées (**Figure IV-26**). Par conséquent, rPhoSL présente un nouveau type de repliement que nous avons appelé β -prisme III.

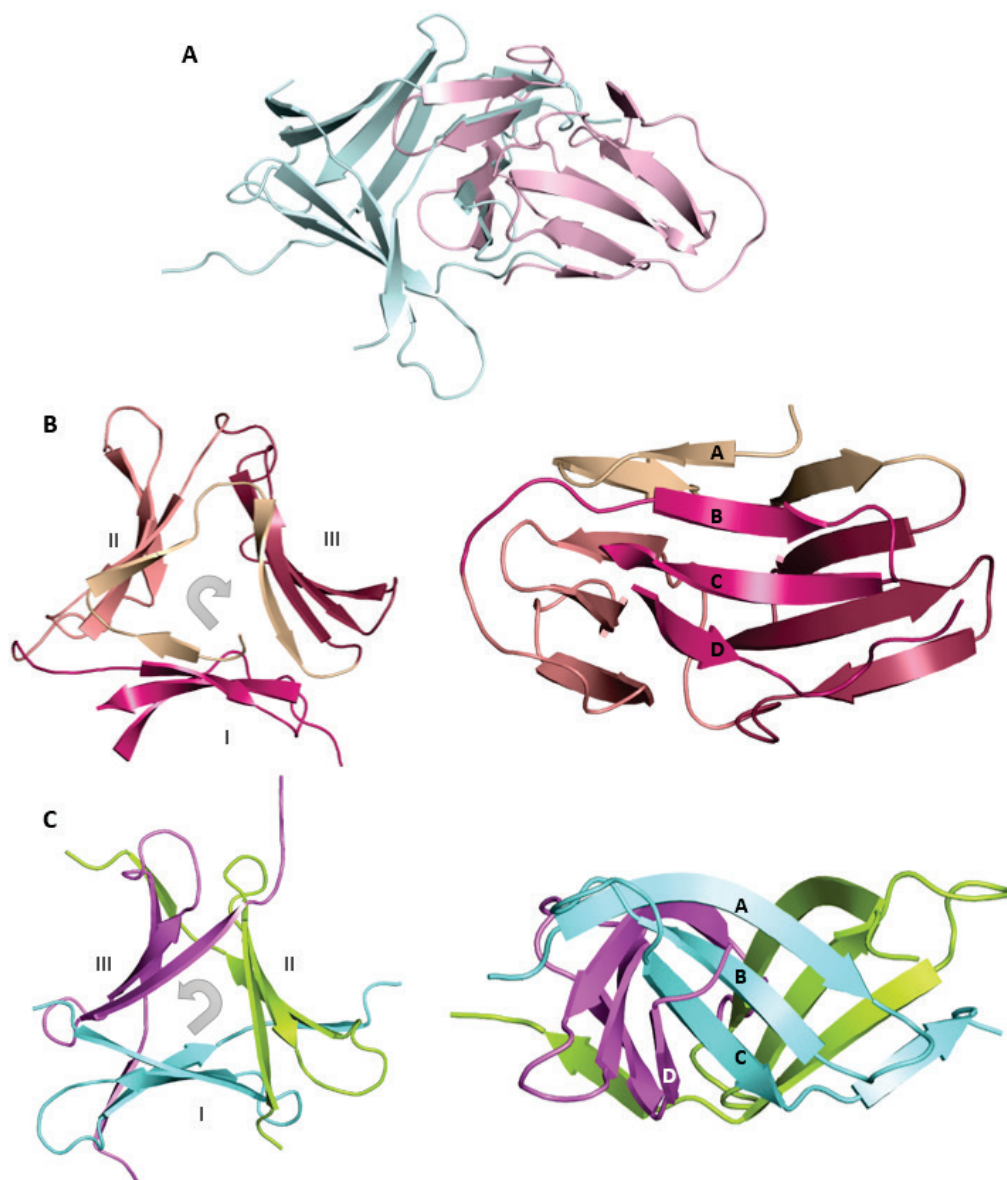


Figure IV-26. Différences structurales entre rPhoSL (β -prisme III) et GNA (β -prisme II).

A. Superposition du β -prisme III de rPhoSL (bleu) avec le β -prisme II de *Galanthus nivalis* agglutinin (GNA, PDB 1MSA, rose). B. Vue de dessus et de côté de GNA (*Galanthus nivalis*) colorée par feuillet β . C. Vue de dessus et de côté de rPhoSL colorée par monomère.

Récemment, la structure du peptide de rPhoSL, obtenue par synthèse chimique, a été résolue par RMN (Yamasaki et al., 2018) et par cristallographie (Yamasaki et al., 2019). Elles

sont similaires à la structure de rPhoSL avec des rmsd respectifs de 1.208 et 0.968 (**Figure IV- 27**).



Figure IV- 27. Superposition des structures de rPhoSL et des structures RMN et cristallographiques du peptide synthétique.

Structure de rPhoSL en vert (pdb 6eke). Structure RMN du peptide chimique en beige (pdb 5xzk). Structure cristallographique du peptide chimique en marron (pdb 6a86).

2.5.3 Résolution de la structure de rPhoSL en complexe avec des N-glycanes core-fucosylés

L'objectif était ensuite d'analyser les structures de rPhoSL en présence de ses ligands afin notamment de comprendre les bases moléculaires de la spécificité de cette lectine pour les N-glycanes fucosylés en α 1,6 de structure variée.

Pour cela, des essais de co-cristallisation ont été réalisés avec rPhoSL et les différents N-glycanes que nous avons à disposition. Finalement, des cristaux de rPhoSL co-cristallisés avec les N-glycanes 1, 2, 3 et 5 (**Figure IV-28**) ont pu être obtenus. Les données ont été collectées au synchrotron. Les structures ont été résolues à 2.1, 1.7, 1.7 et 1.74 Å respectivement par remplacement moléculaire en utilisant les coordonnées de la structure native comme modèle. L'unité asymétrique contient soit quatre trimères de rPhoSL (N-glycane 1, groupe d'espace P2₁22₁), soit deux monomères qui forment les trimères par symétrie cristallographique (N-glycanes 2, 3 et 5, groupe d'espace I2₁3). Les statistiques de collecte et d'affinement sont résumées dans le **Tableau IV- 4**.

	rPhoSL-glycane 1	rPhoSL-glycane 2		rPhoSL-glycane 3		rPhoSL-Glycane 5	
Données collectées							
Ligne de lumière	ESRF ID23-1	SOLEIL Proxima 1		SOLEIL Proxima 1		ESRF BM30 A	
Groupe d'espace	P2 ₁ 22 ₁	I2 ₁ 3		I2 ₁ 3		I2 ₁ 3	
Dimensions maille a b c (Å) α β γ (°)	91.0 95.6 110.6 90 90 90	101.7 101.7 101.7 90 90 90		101.9 101.9 101.9 90 90 90		101.45 101.45 101.45 90 90 90	
Longueur d'onde	0.976	0.979		0.979		0.979	
Resolution (Å)	47.88-2.10 (2.16-2.10)	36.01-1.7 (1.73-1.70)		41.54-1.7 (1.73-1.7)		41.42-1.74 (1.77-1.74)	
Réflexions mesurées/unicues	296329/56993	628467/19497		335255/19409		164813/17987	
Rmerge	0.045 (0.466)	0.048 (0.550)		0.050 (0.663)		0.034 (0.725)	
Mean I / sd(I)	18.6 (3.2)	45.9 (7.3)		32.8 (4.6)		30.3 (3.0)	
Complétude (%)	99.9 (100)	100.0 (100.0)		100.0 (100.0)		100.0 (100.0)	
Multiplicité	5.2 (5.4)	32.2 (29.7)		17.3 (16.9)		9.2 (9.2)	
CC1/2	0.999 (0.907)	1.000 (0.968)		1.000 (0.917)		1.000 (0.824)	
Affinement							
Résolution (Å)	47.88-2.10	36.01-1.70		41.54-1.70		41.42-1.74	
No. of work/ free Reflections	57008 / 2785	19500 / 976		19412 / 973		17992 / 1006	
R _{work} / R _{free} (%)	18.4 / 21.9	16.4 / 18.8		15.7 / 18.7		16.5 / 19.4	
Rmsd bond lengths (Å)	0.015	0.0153		0.0145		0.0157	
Rmsd bond angles (°)	1.84	1.89		2.06		2.04	
Nb d'atomes	Pour les 12 chaînes	Chaîne A	Chaîne B	Chaîne A	Chaîne B	Chaîne A	Chaîne B
Protéine	3801	315	331	317	331	314	331
Ligand	889	113	88	133	112	116	102
Eau	320	145		151		130	
Protéine	40.3	26.4	25.7	24.2	22.9	28.4	28.8
Ligand	50	40.6	33.7	38	31	45.6	43.6
Eau	45.6	40.9		40.4		44.0	
Ramachandran plot (%) Allowed/Favored/Outliers	100 / 97.4 / 0	100 / 98.7 / 0		100 / 98.7 / 0		100 / 98.7 / 0	
code PDB	6FX1	6FX2		6FX3			

Tableau IV- 4. Données de collecte et d'affinement pour les structures de rPhoSL en complexe avec les Glycanes 1, 2 3 et 5.
Les valeurs entre parenthèses correspondent à la couronne de plus haute résolution. Rfree a été calculé avec 5% des réflexions.

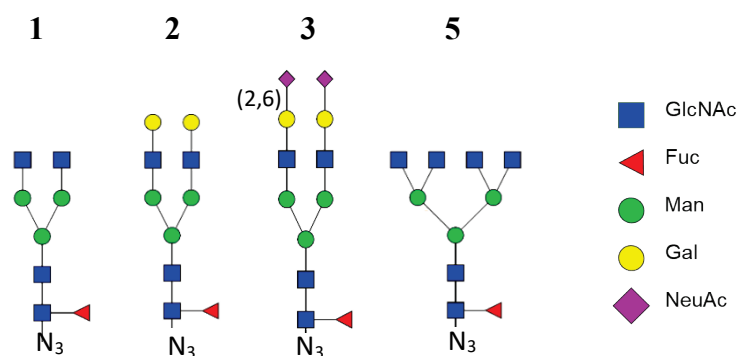


Figure IV-28. Schéma des N-glycanes avec lesquels rPhoSL a été co-cristallisée et dont la structure a été résolue.

La superposition des structures en absence et en présence de N-glycanes montre qu'il n'y a que de légers mouvements de la structure protéique globale. Les N-glycanes ont pu être construits dans la densité électronique sans ambiguïté, à chaque interface entre deux monomères, ce qui confirme le fait que la trimérisation de rPhoSL est nécessaire à son activité (**Figure IV-29**). On remarque cependant une grande flexibilité du N-terminal de rPhoSL, qui se trouve à proximité du site d'interaction. Ce N-terminal est plus long que dans la séquence d'origine, du fait de l'ajout des trois acides aminés GAM après coupure de l'étiquette par la protéase TEV (**Figure IV-16**). Même si cela n'a pas d'impact apparent sur la fixation des sucres dans la structure cristallographique, cela pourrait gêner l'interaction et expliquer la faible affinité mesurée (100 fois inférieure à celle de PhoSL native et de son homologue SL2-1).

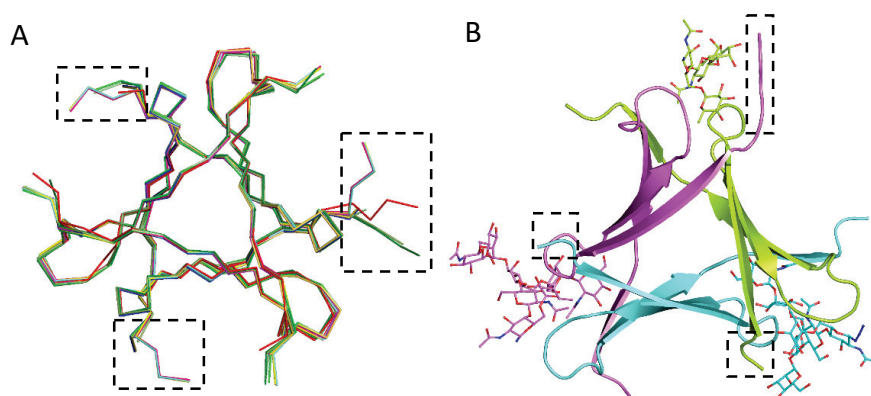


Figure IV-29. A. Superposition des trimères de rPhoSL en absence de ligand avec les structures de rPhoSL en interaction avec les N-glycanes 1, 2, 3 et 5. B. Représentation du N-glycane 1 à chaque interface entre deux monomères.

A. Sans ligand : rouge ; N-glycane 1 : chaînes ABC, DEF, GHI et JKL en nuances de vert. N-glycane 2 : chaînes A et B nuances de jaune ; N-glycane 3 : chaînes A et B en nuances de bleu ; N-glycane 5 : chaînes A et B en nuances de mauve. B. Coloration par monomère. Nter entourés en pointillés.

Site de reconnaissance de la core-fucosylation

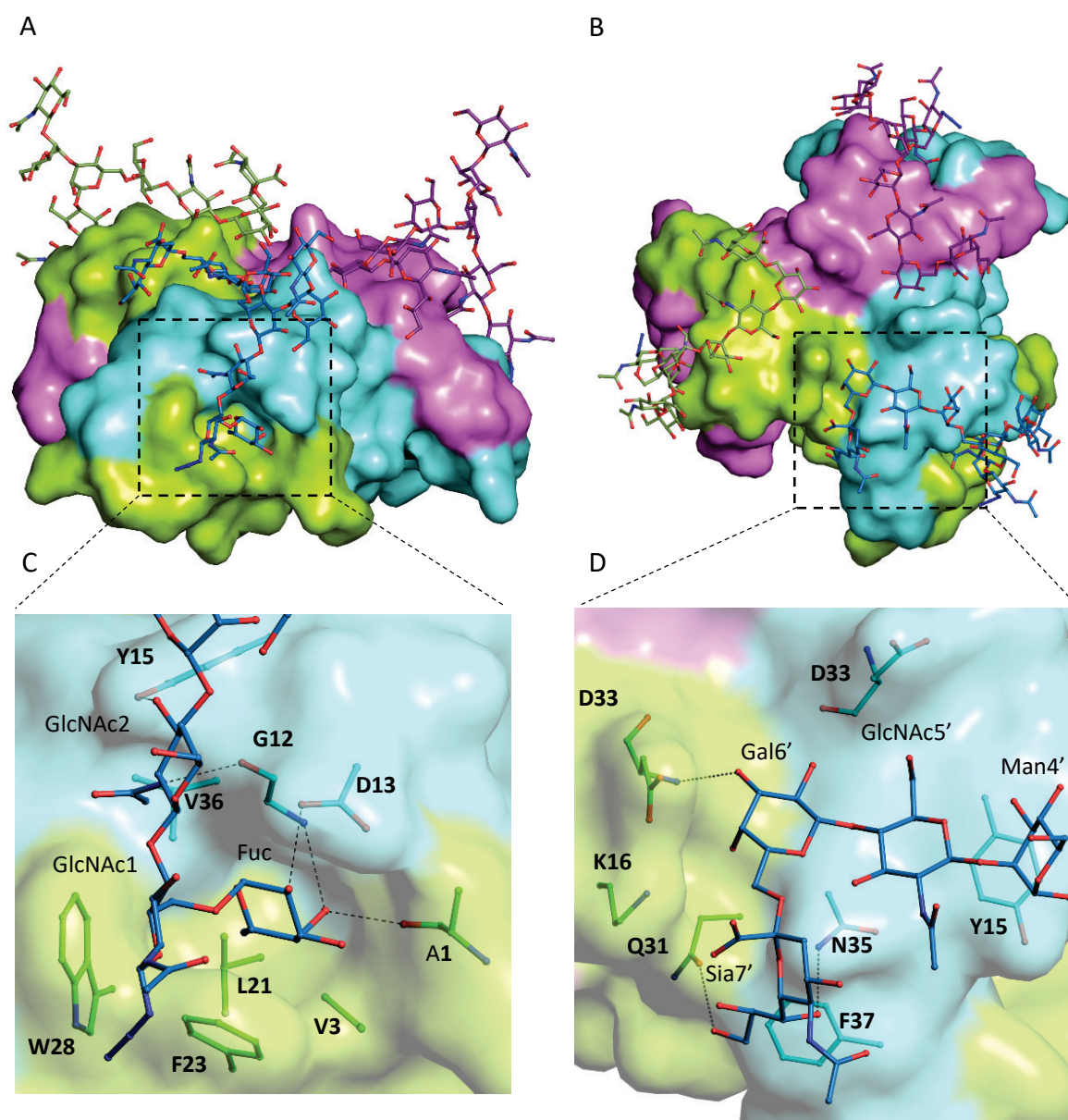


Figure IV-30. Description des interactions entre rPhoSL et le N-glycane 3.

A. Représentation de surface de rPhoSL vue de côté avec ses ligands. B. Vue de dessus. C. Description des interactions avec les 3 premiers monomères du N-glycane 3. D. Description des interactions avec la branche 1,6 du N-glycane 3.

Le fucose et le groupement N-acétyle du GlcNAc 2 se trouvent dans un sillon profond formé essentiellement par les boucles de surface. Un côté de ce sillon est composé d'acides aminés impliqués dans des interactions hydrogènes (alanine 1*, glycine 12, aspartate 13 et tyrosine 15), le fond du sillon est constitué de résidus hydrophobes (valine 3*, leucine 21* et valine 36) et l'autre côté du sillon est constitué de résidus aromatiques (phenylalanine 23* et tryptophane 28) qui génèrent des interactions d'empilement (les « * » indiquent les résidus du monomère voisin, *Figure IV-30*). Le groupement méthyle du fucose est ancré profondément

dans une poche de ce sillon et le cycle du fucose fait des interactions CH- π avec le groupement phényle de la phénylalanine 23. L'architecture du site de fixation et les interactions mises en jeu ont pour conséquence que seul un fucose lié en α 1, 6 peut être reconnu par rPhoSL. Un fucose lié en α 1,2, α 1,3 ou α 1,4 entraînerait des conflits stériques avec la protéine.

Reconnaissance des branches des N-glycanes

Dans toutes les structures, la branche α 1,3 des N-glycanes n'interagit pas avec la protéine et est complètement exposée au solvant. De ce fait, dans la plupart des chaînes protéiques, la densité électronique n'a pas permis de construire cette branche sauf quand elle est stabilisée par des contacts cristallins. À l'opposé, la branche α 1,6 est fortement reconnue et a pu être modélisée dans la densité électronique pour les N-glycanes 2 et 3 et pour huit molécules sur 12 en complexe avec le N-glycane 1 (*Figure VII-1* et *Figure VII-2*). Dans le cas du N-glycane 1, le GlcNAc5' est stabilisé par des interactions CH- π avec le cycle aromatique de la tyrosine 15, et parfois aussi par des interactions hydrogènes avec la chaîne latérale de l'asparagine 35 ou par des contacts avec la tyrosine 15 et la valine 36 via une molécule d'eau (*Figure IV-31.A*). Ces interactions sont cohérentes avec les résultats d'ITC qui montrent une meilleure affinité pour le N-glycane 1 comparé au chitobiose core-fucosylé.

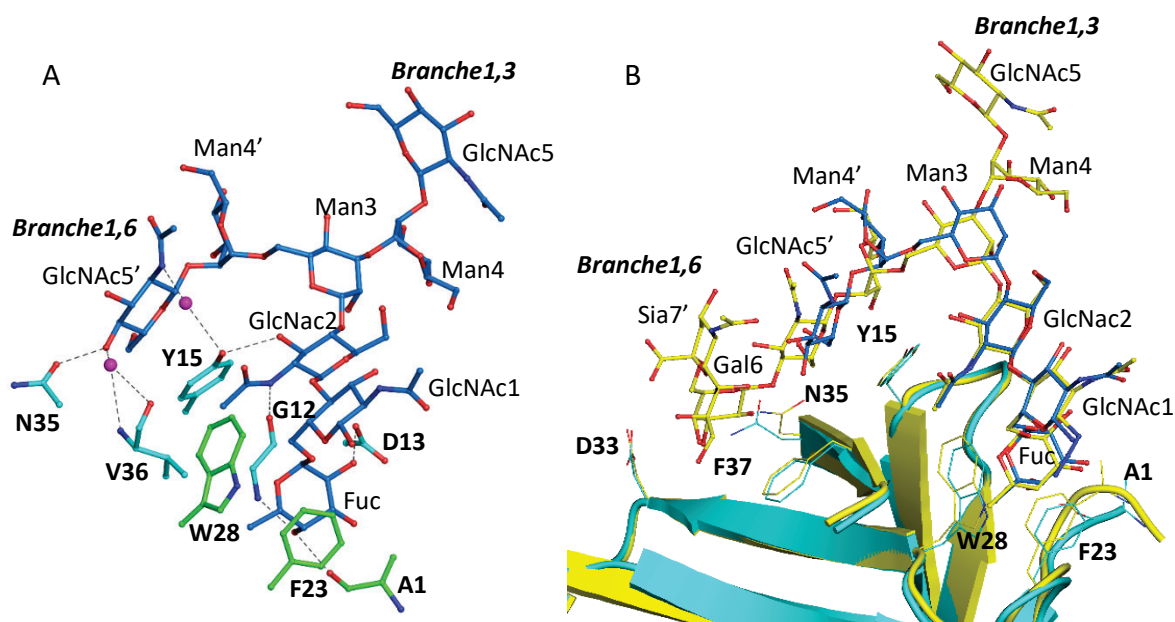


Figure IV-31. A. Description des interactions entre rPhoSL et le N-glycane 1. B. Comparaison des interactions entre le N-glycane 1 (bleu) et le N-glycane 3 (jaune).

A. Résidus en interaction en cyan et vert selon le monomère auquel ils appartiennent. B. rPhosL en interaction avec le N-glycane 1 en cyan et avec le N-glycane 3 en jaune.

Quand le N-glycane est galactosylé (N-glycane 2 ou 3), la branche $\alpha 1,6$ s'étend au niveau d'un deuxième sillon de fixation des sucres (**Figure IV-30**). Celui-ci est formé par la boucle qui connecte les brins C et D et par le brin D d'un monomère ainsi que par la partie C terminale du brin C de l'autre monomère et la chaîne latérale de la lysine 16*. L'azote de la chaîne principale de l'aspartate 33 fait une interaction hydrogène avec le OH-3 du Gal6'. L'acide sialique (Sia7') est dirigé vers le solvant et interagit avec la protéine principalement via des interactions hydrogènes avec la chaîne latérale de la glutamine 31 et de l'asparagine 35 et des interactions hydrophobes avec le groupement phényle de la phénylalanine 37 (**Figure IV-30**).

Globalement, rPhoSL est assez statique et la fixation des sucres influence essentiellement les conformations des boucles BC et la chaîne latérale de l'asparagine 35. Celle-ci peut adopter deux conformations pour le N-glycane 1 mais une seule pour les N-glycane 2 et 3, de façon à éviter les clashes stériques avec le galactose et l'acide sialique (**Figure IV-31.B**). Les mesures des angles diédraux des liaisons glycosidiques ont montré qu'elles correspondaient à des conformations de faible énergie. L'empilement cristallin crée un environnement légèrement différent pour chaque site d'interaction, ce qui a permis de rendre compte de la flexibilité des N-glycane (**Figure IV-32**).

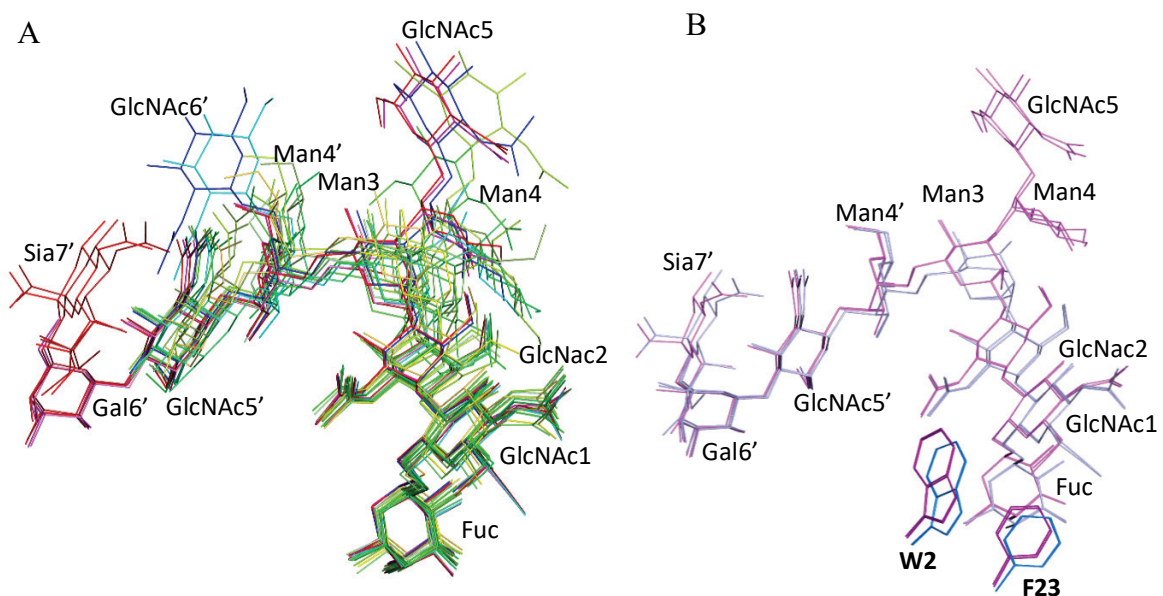


Figure IV-32. Flexibilité des N-glycane en interaction avec rPhoSL dans les structures cristallographiques.

A. Superposition du N-glycane 1 (12 chaînes en nuances de verts), N-glycane 2 (chaînes A et B en nuances de violet), N-glycane 3 (chaînes A et B en nuances de rouge), N-glycane 5 (chaînes A et B en nuances de bleu). B. Superposition des N-glycane 2 et 3 (chaînes A en violet, chaînes B en bleu-pâle) avec W28 et F23 correspondants en violet et bleu foncés.

À cause des contraintes imposées par le site de fixation du fucose, la mobilité du fucose et du groupement N-acétyle du GlcNAc2 est assez limitée. De petites différences d'angles dièdres de la liaison α 1,6 sont à l'origine de légers décalages au niveau du chitobiose, en particulier du GlcNAc 2. Mais cela s'accompagne de légers changements d'orientation de la phénylalanine 23 et du tryptophane 28, ce qui permet d'optimiser les interactions hydrophobes (**Figure IV-32.B**). Une grande mobilité est observée pour la branche α 1,3 qui est complètement exposée au solvant. Les mouvements de la branche α 1,6 sont beaucoup plus restreints, en particulier lorsqu'on est en présence de galactose qui interagit avec la protéine.

Concernant le N-glycane 5, seuls 9 ou 8 monomères sur les 10 ont pu être construits en densité pour les chaînes A et B respectivement (**Figure VII-2**). Cependant, les interactions avec rPhoSL sont les mêmes que pour le N-glycane 1 car les deux branchements supplémentaires constitués du GlcNAc6 (branche 1,3) et du GlcNAc6' (branche 1,6) n'interagissent pas avec la protéine (**Figure IV-33**). À cause de cette absence d'interaction, les monomères de la branche 1,3 et le GlcNAc 6' sont peu stabilisés et leur construction dans la densité électronique est difficile. La branche 1,6 est principalement stabilisée par l'interaction CH- π avec le cycle de la tyrosine 15.

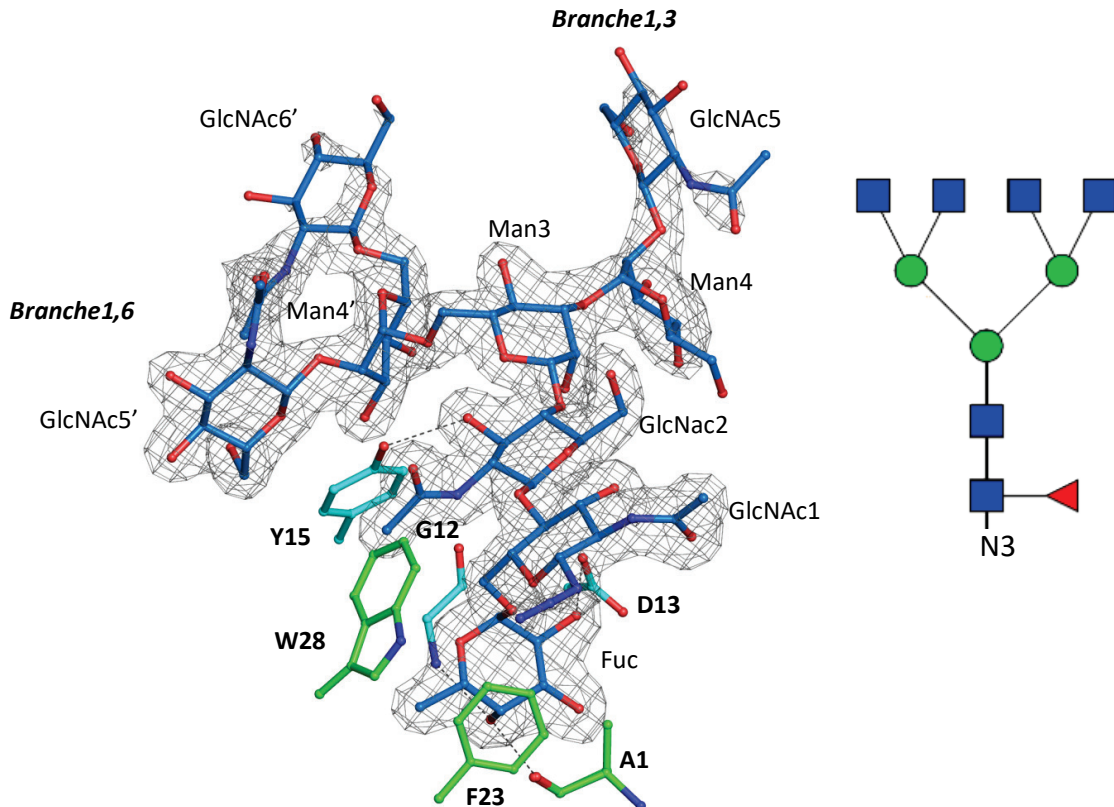


Figure IV-33. Description des interactions de rPhoSL avec le N-glycane 5 en densité. Coloration des acides aminés par monomère. Représentation schématique à droite.

2.6 Définition d'une nouvelle famille de lectines

La recherche d'homologues à partir de la séquence de rPhoSL a révélé que la majorité des séquences homologues sont composées de plusieurs répétitions de la séquence de rPhoSL, comme c'est le cas pour les protéines rPhoNL et SL2-1, à part pour RSL qui n'est constituée que d'une seule répétition (**Figure IV-34**). L'alignement de séquence montre la stricte conservation de deux cystéines formant un pont disulfure ainsi que de la plupart des résidus impliqués dans les interactions avec le fucose (G12, L21, F23, W28). Une séquence consensus pour les protéines de type PhoSL a ainsi été proposée : ${}_{10}\text{CDGDX YXCXAXLXFGDGXW}_{28}$ (X représentant n'importe quel acide aminé).

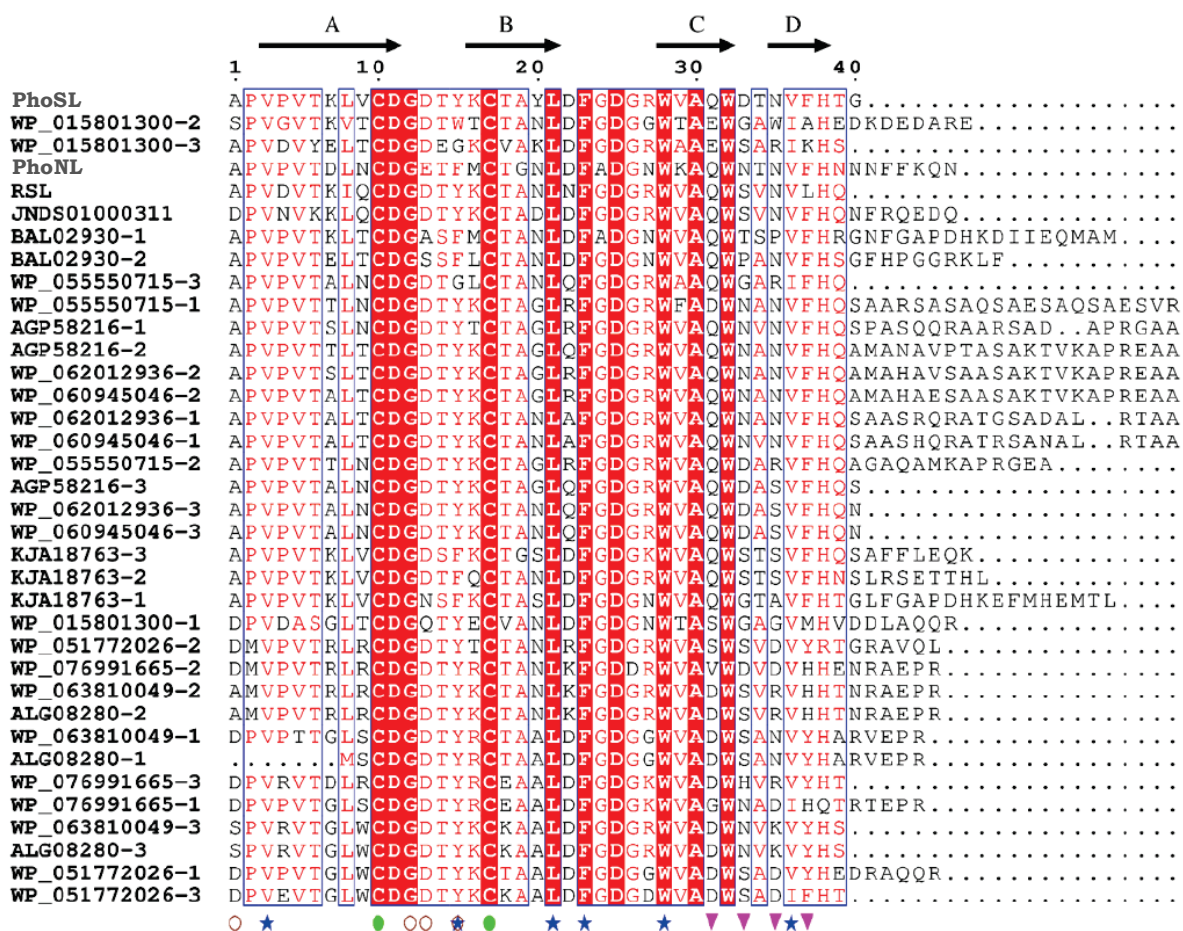


Figure IV-34. Alignement multiple de rPhoSL.

Les séquences alignées sont issues de *Pholiota squarrosa* (PhoSL); *Actinosynnema mirum* (WP_015801300); *Pholiota nameko* (PhoNL); *Rhizopus delemar* Type II NRRL 21477 (RSL, NODE_1316_length_1457_cov_6.197667_1), *Rhizopus stolonifer* B9770 (JNDS01000311, shotgun sequence ctg7180000282406_1); *Streptomyces* sp. NBRC 110028 (WP_055550715); *Streptomyces sporocinereus* (WP_062012936); *Streptomyces hygroscopicus* (WP_060945046) *Streptomyces rapamycinicus* NRRL 5491 (AGP58216); *Hypholoma sublateritium* FD-334 SS-4 (KJA18763); *Saccharothrix* sp. NRRL B-16314 (WP_051772026); *Actinosynnema* sp. ALI-1.44 (WP_076991665); *Kibdelosporangium*

phytohabitans (WP_063810049 and ALG08280). Les peptides signaux ont été supprimés pour simplifier la lecture. La numérotation de PhoSL et le nom des différents brins sont indiqués au-dessus. Les cystéines impliquées dans les ponts disulfures sont indiquées par un cercle vert, les résidus impliqués dans des interactions hydrophobes ou des interactions hydrogènes avec le chitobiose core-fucosylé sont indiqués par une étoile bleue et un cercle marron respectivement et les résidus qui interagissent avec la branche 1,6 sont indiqués par un triangle. La figure a été réalisée avec ESPript 3.0.

Cette séquence consensus définit une nouvelle famille de lectines et peut être utilisée pour prédire de nouveaux membres et la capacité d'interaction avec le fucose α 1,6. Les variations observées pour l'aspartate 13 et la tyrosine 15, par exemple chez PhoNL, pourraient entraîner une baisse d'affinité en impactant les interactions hydrogènes. Les mutations de la valine 36 en isoleucine 36 ne devraient pas avoir beaucoup d'impact sur l'interaction avec les ligands puisque le fucose pourrait toujours s'insérer. La conservation de la séquence est plus faible pour le second site d'interaction, ce qui peut impacter la fixation de la branche α 1,6. Les neuf résidus au niveau du C-terminal des répétitions semblent donc responsables de la spécificité fine des protéines de type PhoSL.

2.7 Mutations de rPhoSL (rPhoSLm), purification et analyse

L'analyse structurale de rPhoSL a permis de mettre en évidence une origine possible des faibles affinités mesurées en ITC pour les N-glycanes comparé aux valeurs obtenues pour la protéine native et pour SL2-1. Il apparaît en effet dans la structure que les trois acides aminés en N-terminal « GAM », qui ne sont pas présents dans la séquence d'origine, et qui proviennent de la coupure de l'étiquette avec la protéase TEV, sont très proches du site d'interaction. Bien qu'ils ne semblent pas empêcher la fixation des sucres sur les trois sites d'interaction, ils peuvent, de par leur nature désordonnée, gêner l'insertion des N-glycanes et provoquer une diminution de l'affinité. L'objectif a donc été de muter ce N-terminal de façon à diminuer au maximum les acides aminés ajoutés en N-terminal sur la protéine purifiée. La protéine mutante a ensuite été analysée pour évaluer l'impact de la suppression des acides aminés ciblés.

Seuls les deux acides aminés alanine et méthionine ont été supprimés car la glycine fait partie de la séquence consensus de coupure de la protéase TEV et est donc indispensable pour pouvoir couper l'étiquette N-terminale. La suppression de ces deux acides aminés a été réalisée par mutagenèse dirigée. La lectine ainsi mutée (rPhoSLm) a ensuite été produite et purifiée selon le même protocole que précédemment, et a été obtenue avec des rendements similaires.

L'analyse de sa stabilité thermique par TSA indique que rPhoS_Lm est plus stable qu'avant mutation puisque aucune dénaturation n'a été observée entre 0 et 100°C alors qu'auparavant rPhoS_L avait un T_m de 68°C. Le mutant était toujours capable d'hémagglutiner les globules rouges de lapin. Quand les globules rouges sont préalablement traités à la trypsine, rPhoS_Lm a une capacité d'hémagglutination 2 fois meilleure à celle de la protéine non mutée (*Figure IV-35*). Il n'y a pas de différence en absence de traitement. Cela semble indiquer une meilleure reconnaissance des ligands peu accessibles en surface par la protéine mutée, ce qui pourrait être associé à une meilleure affinité.

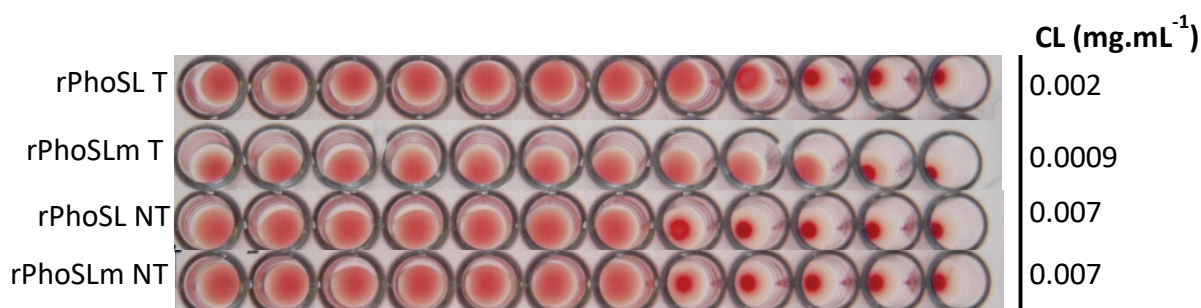


Figure IV-35. Comparaison des capacités d'hémagglutination des globules rouges de lapin entre rPhoS_Lm et rPhoS_L.

Les globules rouges ont été traités (T) ou non (NT) à la trypsine avant le test. CL = concentration limite de lectine pour laquelle il y a encore hémagglutination. rPhoS_L est à 0.45 mg.mL⁻¹ et rPhoS_Lm est à 0.5 mg.mL⁻¹ dans le premier puits.

Ces premiers résultats ont donc montré que le mutant produit était stable et actif. L'étude a donc été poursuivie afin de déterminer si les mutations réalisées au niveau du N-ter de la protéine avaient effectivement un impact sur l'affinité de rPhoS_L par ITC.

L'affinité pour le chitobiose core-fucosylé est environ dix fois meilleure après mutation avec un K_D de 20 μ M, ce qui confirme que les acides aminés ajoutés en N-terminal ont bien un impact défavorable sur l'affinité de la lectine (*Tableau IV-5*). Cette affinité, autour de 20 μ M est tout à fait correcte pour l'utilisation de rPhoS_Lm en tant qu'outil de marquage, même si on peut remarquer qu'il existe toujours un facteur 10 comparé aux valeurs d'affinité de la lectine native et de SL2-1. Cela suggère que la glycine qui a été conservée en N-ter en raison de son implication dans la coupure de l'étiquette par la protéase TEV pourrait encore impacter négativement l'affinité de rPhoS_Lm.



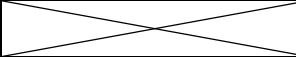


N-glycanes	rPhoSL	rPhoSLm
	Affinité (μM)	
Chitobiose core-fucosylé 	180	20
N-glycane 1 	105	
N-glycane 3 		28

Tableau IV-5 Comparaison des affinités de rPhoSL et rPhoSLm mesurées par ITC.

rPhoSL et rPhoSLm à 50 μM ont été titrées avec 5 mM de chitobiose core-fucosylé et 10 mM du N-glycane 1 ou 3 à 25°C. La protéine et les sucres ont été préparés dans 25 mM MES pH6. Les mesures ont été réalisées deux fois.

De plus, nous avons vu dans le paragraphe précédent que l'acide sialique et le galactose de la branche $\alpha 1, 6$ du N-glycane 3 interagissent au niveau d'un deuxième sillon de rPhoSL différent de celui du fucose. L'affinité pour ce sucre a donc été mesurée pour évaluer l'impact de ces interactions supplémentaires sur la force de l'interaction, mais les résultats montrent une affinité de même ordre de grandeur à celle mesurée pour le chitobiose core-fucosylé (28 μM). Le fait qu'on ne mesure pas d'augmentation d'affinité peut s'expliquer par une perte de flexibilité du ligand. En effet, la restriction des mouvements d'un ligand cause une perte d'entropie conformationnelle, définie par le nombre de conformations possibles d'une molécule, ce qui est défavorable à l'interaction.

rPhoSLm est un outil à fort potentiel dans la reconnaissance spécifique de la core-fucosylation. Cette lectine peut être produite sous forme recombinante, ce qui est un gage de qualité et permet d'envisager son utilisation dans diverses applications, notamment en diagnostic des cancers. Ce travail a permis de comprendre les détails de l'interaction au niveau structural, de définir une séquence consensus pour la reconnaissance de la core-fucosylation et de révéler un nouveau type de repliement protéique en β -prisme III. Cela va permettre de trouver plus facilement de nouvelles lectines avec ce type de spécificité et, par homologie, d'améliorer les capacités des lectines existantes, comme cela a été réalisé ici avec le mutant rPhoSLm.

Chapitre 3.

**Reconnaissance de la core-
fucosylation : rPhoNL**

3. Reconnaissance de la core-fucosylation : rPhoNL

Les recherches de séquences homologues à PhoSL ont permis de révéler l'existence d'une autre lectine fongique issue d'un autre champignon de genre *Pholiota* : *Pholiota nameko*, que l'on a appelé PhoNL. PhoNL est une protéine de 16.58 kDa qui présente trois répétitions de séquences homologues à celle de PhoSL avec en moyenne 70 % d'identité de séquence. L'alignement des séquences de ces deux lectines est présenté **Figure IV-36**. L'alignement multiple a été réalisé avec Clustal Omega (Chojnacki et al., 2017) puis mis en forme avec ESPript3 (Robert and Gouet, 2014).

	1	10	20	30	40	50									
rPhoNL1	APVPVT	KLTC	CDGA	SFMC	TANL	DFAD	DGNW	VQAQ	WTSP	VFHR	GNFG	APD	HKDI	IEQM	MAM
rPhoNL2	APVPVT	ELTC	CDGS	SFLC	TANL	DFGD	DGNW	VQAQ	WPAN	VFHS	GFHP	GGR	..	KL	F.....
rPhoNL3	APVPVT	DLNC	CDGE	TFMC	TGNL	DFAD	DGNW	KAQW	NNTN	VFNH	NNNF	FK	QN
rPhoS _L	APVPVT	KLVC	CDGD	TYKC	TAYL	DFGD	DGRW	VQAQ	WDTN	VFHT	GNFG	APD	HKDI	IEQM	MAM

Figure IV-36. Alignement des séquences de rPhoSL et de rPhoNL.

rPhoNL 1, 2 et 3 correspondent aux trois répétitions de la séquence de rPhoSL.

Ces homologies permettent de penser que PhoNL pourrait avoir des propriétés similaires à PhoSL pour la reconnaissance spécifique de la core-fucosylation. PhoNL pourrait donc avoir le potentiel de devenir comme PhoSL un outil de reconnaissance spécifique pour ce type de motif. Le fait que les trois répétitions soient présentes au sein d'un même monomère pourrait également apporter des propriétés intéressantes et pourrait peut-être améliorer les rendements d'expression dans *E. coli*. De plus, les différences entre les séquences de PhoSL et PhoNL pourraient permettre de mieux évaluer l'implication de chaque acide aminé dans les interactions avec les N-glycanes. Cette lectine a donc été produite également sous forme recombinante, en parallèle de la production de rPhoSL avant d'être analysée.

3.1 Production et purification de rPhoNL

Le gène d'origine de PhoNL contient une séquence codant pour un peptide signal en N-terminal, ce qui indique qu'il s'agit d'une protéine sécrétée dans le champignon. Cette séquence a par la suite été coupée lors du clonage car elle causait de l'instabilité sur rPhoNL. Le prédicteur Disulfind (Ceroni et al., 2006) a indiqué la présence de ponts disulfures, comme pour rPhoSL. De ce fait, le mode de production choisi a été le même que celui de rPhoSL : la lectine a été clonée dans le vecteur pET-39-TEV afin de permettre une expression dans le périplasme de souches BL21 Star (DE3) d'*E. coli*, puis elle a été purifiée selon un protocole identique. Les résultats des étapes de purification sur colonne IMAC puis DEAE sont présentés dans la **Figure IV-37**. Ils montrent que rPhoNL a pu être correctement purifiée à

environ 95 %, ce qui a permis ensuite de réaliser sa caractérisation biochimique et structurale. Elle a été concentrée dans 25 mM MES pH 6 et conservée à 4°C. Le rendement de production et de purification est d'environ 2 mg par litre de culture, et est donc supérieur à celui de rPhoSL (1mg par litre).

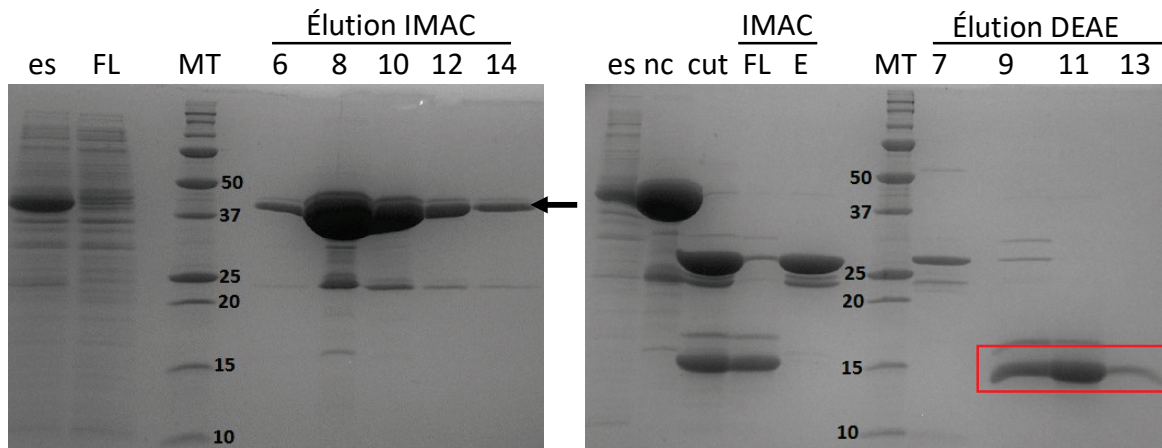


Figure IV-37. Gels SDS-PAGE des deux étapes de purification de rPhoNL sur IMAC et DEAE.

Gels SDS-PAGE Tris glycine 15%. es = extrait soluble ; FL = fraction non retenue ; MT = marqueur de taille en kDa ; E : élution sur colonne IMAC. La flèche indique la taille attendue pour rPhoNL-étiquette.

Les modes de production et purification étant identiques à ceux de rPhoSL, là aussi, trois acides aminés « GAM » ont été ajoutés en N-ter de la protéine suite au clonage et à la coupure de l'étiquette par la protéase TEV. En effet, l'impact de ces trois acides aminés n'était pas encore connu au moment de la purification de rPhoNL.

3.2 Analyse de l'état d'oligomérisation, de la stabilité et de l'activité

Des analyses de l'état d'oligomérisation de rPhoNL ont été réalisées en amont de ce travail par gel filtration. Les résultats semblaient montrer que rPhoNL formait des monomères, voire des dimères dans certaines conditions, mais le manque de reproductibilité de ces expériences n'a pas permis de conclure quant à la capacité de dimérisation de cette lectine. Des expériences complémentaires ont été réalisées par la suite pour essayer de comprendre si cette lectine est capable de dimériser ou non, et seront décrites plus loin dans ce chapitre.

Ensuite, l'activité et la stabilité de rPhoNL ont été étudiées afin de vérifier la qualité de la protéine purifiée. L'activité de rPhoNL a été évaluée par un test d'hémagglutination, qui permet de vérifier rapidement et avec peu de matériel si rPhoNL est capable de se fixer à ses ligands, et si elle est multimérique. rPhoNL est capable d'hémagglutiner les globules rouges,

elle est donc active et multivalente (**Figure IV-38**). Sa capacité d'hémagglutination est environ 4 fois inférieure lorsque les globules rouges sont préalablement traités à la trypsine. Ce résultat est étonnant car l'effet inverse avait été observé pour rPhoSL qui a une meilleure capacité d'hémagglutination des globules rouges traités à la trypsine. Cela laisse supposer que rPhoNL n'a pas accès aux mêmes ligands que rPhoSL, ce qui peut s'expliquer à ce stade de l'étude par la reconnaissance de ligands différents, des affinités différentes, ou des structures protéiques qui ne permettent pas la même accessibilité. rPhoNL a une capacité d'hémagglutination globalement inférieure à rPhoSL quelle que soit la condition de traitement à la trypsine, avec une concentration limite d'hémagglutination 500 fois plus faible après traitement et environ 35 fois plus faible sans traitement. rPhoNL semble donc avoir une moins bonne capacité de reconnaissance de ses ligands que rPhoSL.

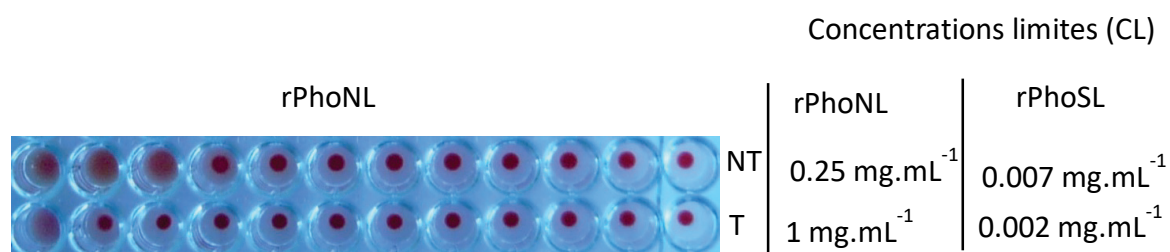


Figure IV-38. Test d'hémagglutination sur globules rouges de lapin de rPhoNL pour du sang préalablement traité (T) ou non (NT) à la trypsine. Comparaison avec rPhoSL.
CL = concentration limite de lectine pour laquelle il y a encore hémagglutination.

La stabilité de rPhoNL a ensuite été étudiée au cours de tests de TSA (Thermal Shift Assay) afin de déterminer le tampon dans lequel la lectine est la plus stable. Les tests ont montré que rPhoNL présentait des dénaturations en deux étapes, en particulier pour des pH acides. Elle a une température de fusion toujours supérieure à 50°C avec un maximum à 61°C pour les dénaturations en une étape et un maximum vers 80°C pour les dénaturations en deux étapes (**Figure IV-39**). Les deux étapes de dénaturation peuvent s'expliquer par des différences de stabilité au sein des monomères de rPhoNL avec des zones plus sensibles à la température que d'autres. Ou bien, dans l'hypothèse d'une dimérisation de rPhoNL, cela pourrait correspondre à la séparation du dimère, avant la dénaturation de chaque monomère. Cependant, rPhoNL qui a été conservée dans un tampon 25 mM MES pH 6, n'a présenté aucun signe de dénaturation après plusieurs mois à 4°C, ce qui montre que cette protéine est très stable dans le temps et ne pose aucun problème de conservation.

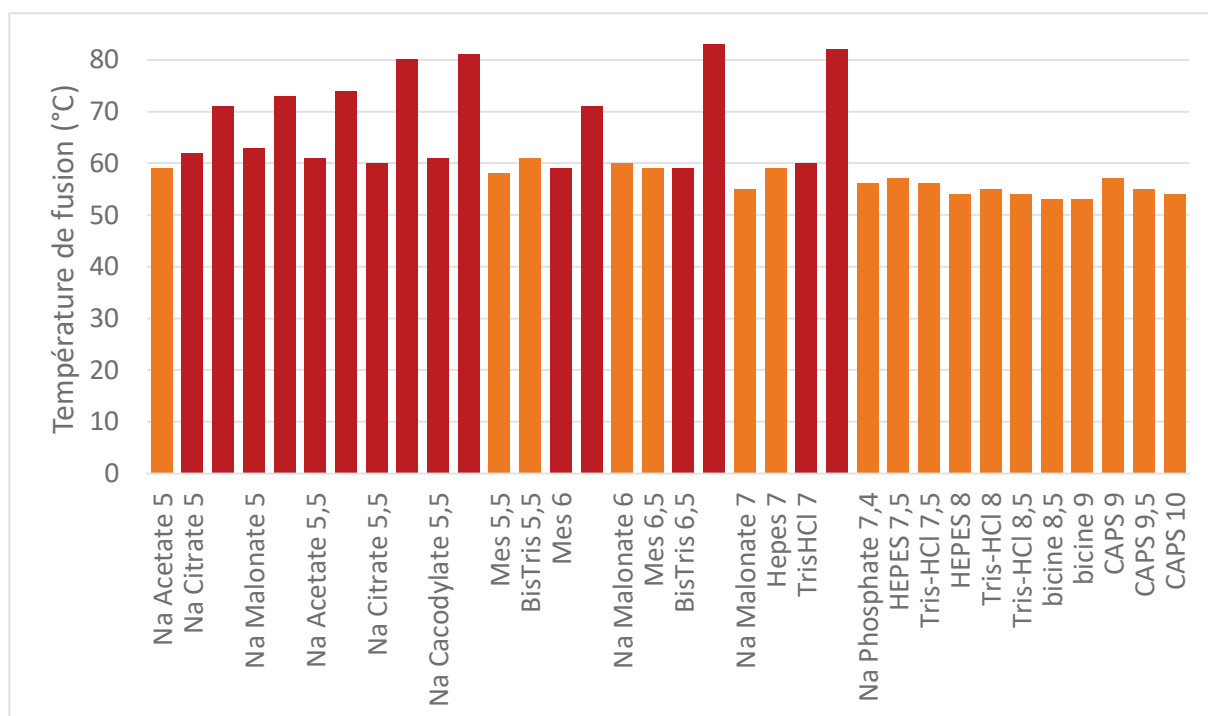


Figure IV-39. Tests TSA avec rPhoNL et 26 tampons de pH 5 à pH 10.

Les tampons ont été dilués à 0.1M. rPhoNL a été utilisée à 0.1 mg.mL⁻¹ dans du sypro orange 5x. Les mesures ont été réalisées sur un appareil de QRT-PCR (Mini-Opticon Biorad). En rouge, les conditions qui conduisent à une dénaturation en deux étapes. L'expérience a été réalisée en duplicats.

3.3 Étude de l'interaction lectine-sucre

3.3.1 Analyse de la spécificité sur puces à sucres (glycan array)

La spécificité de rPhoNL a ensuite été étudiée pour déterminer si cette lectine est également strictement spécifique de la core-fucosylation des N-glycanes. Comme pour rPhoSL, l'analyse a été réalisée sur une puce constituée de 600 sucres présents à la surface des cellules de mammifères (Consortium for Functional Glycomics, CFG). Les résultats montrent que rPhoNL se fixe uniquement aux N-glycanes pour lesquels la core-fucosylation est présente, ce qui confirme sa stricte spécificité pour ce motif (**Figure IV-40**). De plus, le fait que la capacité d'hémagglutination de rPhoNL soit plus faible que celle de rPhoSL n'est donc pas lié à la reconnaissance de ligands différents, mais éventuellement à des affinités différentes ou des structures protéiques qui ne permettent pas la même accessibilité aux ligands sur les globules rouges.

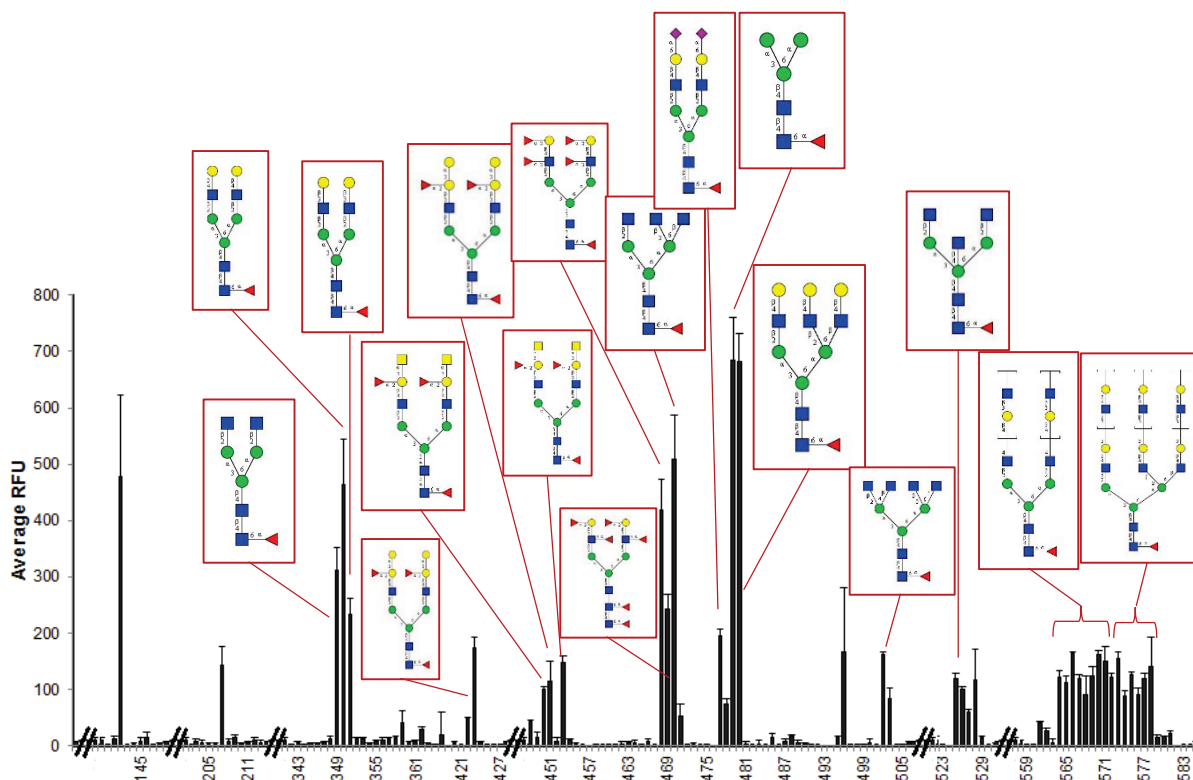


Figure IV-40. Résultats de l'étude de rPhoNL sur puce à sucre version 5.2. rPhoNL ($50 \mu\text{g.mL}^{-1}$) a été marquée à l'alexa 488. Seuls les sucres avec un signal RFU supérieur à 20 % du signal maximal ont été représentés.

3.3.2 Mesure de l'affinité par Titration Calorimétrique Isotherme

L'objectif a ensuite été d'évaluer plus précisément les capacités de reconnaissance de cette lectine en mesurant son affinité pour des N-glycanes core-fucosylés. Le but était aussi de comparer son affinité avec celle de rPhoSL. Des mesures ont été réalisées avec du chitobiose core-fucosylé (N2F N-glycan, carbosynth) ou du N-glycane 1 (octasaccharide). Malheureusement, bien qu'il y ait interaction entre la lectine et le sucre, aucune valeur d'affinité n'a pu être déterminé par ITC car l'affinité était trop faible pour obtenir une courbe correcte. Cela montre donc que rPhoNL a une affinité beaucoup plus faible que rPhoSL pour les sucres core-fucosylés, ce qui est cohérent avec les résultats des tests d'hémagglutination. Il serait intéressant de comprendre d'où vient cette différence d'affinité entre les deux lectines, comme par exemple les acides aminés impliqués ou les déterminants structuraux.

3.1 Étude structurale

En vue de comparaison avec la structure de rPhoSL et pour mieux comprendre les différences d'affinité de rPhoNL à l'échelle atomique, sa structure a été déterminée par cristallographie aux rayons X. L'analyse structurale avait aussi pour but de guider l'optimisation de rPhoNL pour le choix des mutations ciblées par exemple. Pour cette étude,

la lectine a été produite en amont avec des méthionones sélénées (PhoNLSeMet) pour permettre par la suite de déterminer les phases par méthode SAD.

3.1.1 Description de la structure de rPhoNL

Afin d'obtenir des cristaux de rPhoNL, le criblage de conditions de cristallogénèse a été réalisé au laboratoire ou à l'aide d'un robot de cristallisation (HTX, EMBL, Grenoble). La lectine a été utilisée à une concentration de 6.3 mg.mL^{-1} . Les essais de cristallisation en absence de ligand se sont révélés infructueux pour rPhoNL, mais des cristaux ont été obtenus en co-cristallisation avec le N-glycane 1 dans une solution composée de 100 mM MgCl_2 et 5% PEG 8K, 5% PEG 1K (*Figure IV-41*).

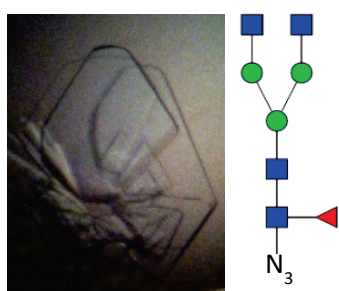


Figure IV-41. Photo de cristaux de rPhoNL co-cristallisée avec le N-glycane 1.

La structure de rPhoNLSeMet, co-cristallisée avec le N-glycane 1 a été résolue à 2.0 \AA par méthode SAD en utilisant le signal anomal du sélénium dans le groupe d'espace $P2_1$ (*Figure IV-42*). Les informations sur la collecte de données ainsi que sur les statistiques de collecte et d'affinement sont indiquées dans le **Tableau IV-6**. rPhoNL forme des dimères dans l'unité asymétrique, les deux monomères étant reliés entre eux par des interactions hydrophobes et deux interactions hydrogènes entre l'histidine 47 d'une chaîne et l'asparagine 138 de l'autre chaîne. Cela pourrait indiquer la capacité de rPhoNL à former des dimères en solution bien qu'il puisse s'agir d'un artefact dû à la cristallisation. La superposition de la structure de rPhoNL avec la structure de rPhoSL montre que rPhoNL a aussi un repliement de type β -prisme III (*Figure IV-42. A et C*). On retrouve les trois échanges de brin entre le brin D des trois répétitions de rPhoNL et la structure est stabilisée par les mêmes types d'interactions que rPhoSL. La différence structurale majeure repose sur la présence de deux grandes boucles de surface (a et b) chez PhoNL qui relient les trois répétitions équivalentes à PhoSL. Le rmsd calculé entre les deux structures, lorsque les atomes de ces boucles ne sont pas pris en compte, est de 0.949, ce qui montre l'homologie structurale de la structure commune aux deux lectines. La boucle **a** n'a pas pu être construite entièrement car la densité n'est pas assez bien

définie pour les acides aminés 54 à 56, du fait de sa flexibilité et de sa forte exposition au solvant.

		rPhoNLSeMet-glycane 1	
Données collectées			
Ligne de lumière		ESRF BM30A	
Groupe d'espace		P2 ₁	
Dimensions de maille a b c (Å) α β γ (°)		47.12 73.75 61.65 90 103.54 90	
Longueur d'onde		0.979	
Résolution (Å)		33.91-2.00 (2.05-2.00)	
Measured reflections		102446	
Unique reflections		27635	
Rmerge		0.065 (0.358)	
Mean I / sd(I)		11.9 (3.2)	
Complétude (%)		99.4 (100)	
Multiplicité		3.7 (3.7)	
CC _{1/2}		90.95(0.865)	
Affinement			
Résolution (Å)		38.91-2.00	
No. of work/ free Reflections		27800 / 1390	
R _{work} / R _{free} (%)		19.4 / 22.7	
Rmsd bond lengths (Å)		0.0158	
Rmsd bond angles (°)		1.75	
Nb d'atomes		Chaîne A	Chaîne B
Protéine		1101	1031
Ligand		50	50
Autres hétéroatomes			1
Eau		261	
B-facteurs moyens			
Protéine		30.16	28.97
Ligand		45.70	41.76
Eau		34.40	
Ramachandran plot (%)		Allowed	100
		Favored	94.8
		Outliers	0

Tableau IV-6. Données de collecte et d'affinement pour la structure de rPhoNLSeMet co-cristallisée avec le N-glycane 1.

Les valeurs entre parenthèses correspondent à la couronne de plus haute résolution. R_{free} a été calculé avec 5% des réflexions.

Ces résultats confirment donc que rPhoNL a une structure homologue à celle de rPhoSL. La présence des grandes boucles de surface reliant les trois répétitions entre elles peut éventuellement avoir un impact sur l'accessibilité de la lectine à des ligands peu accessibles.

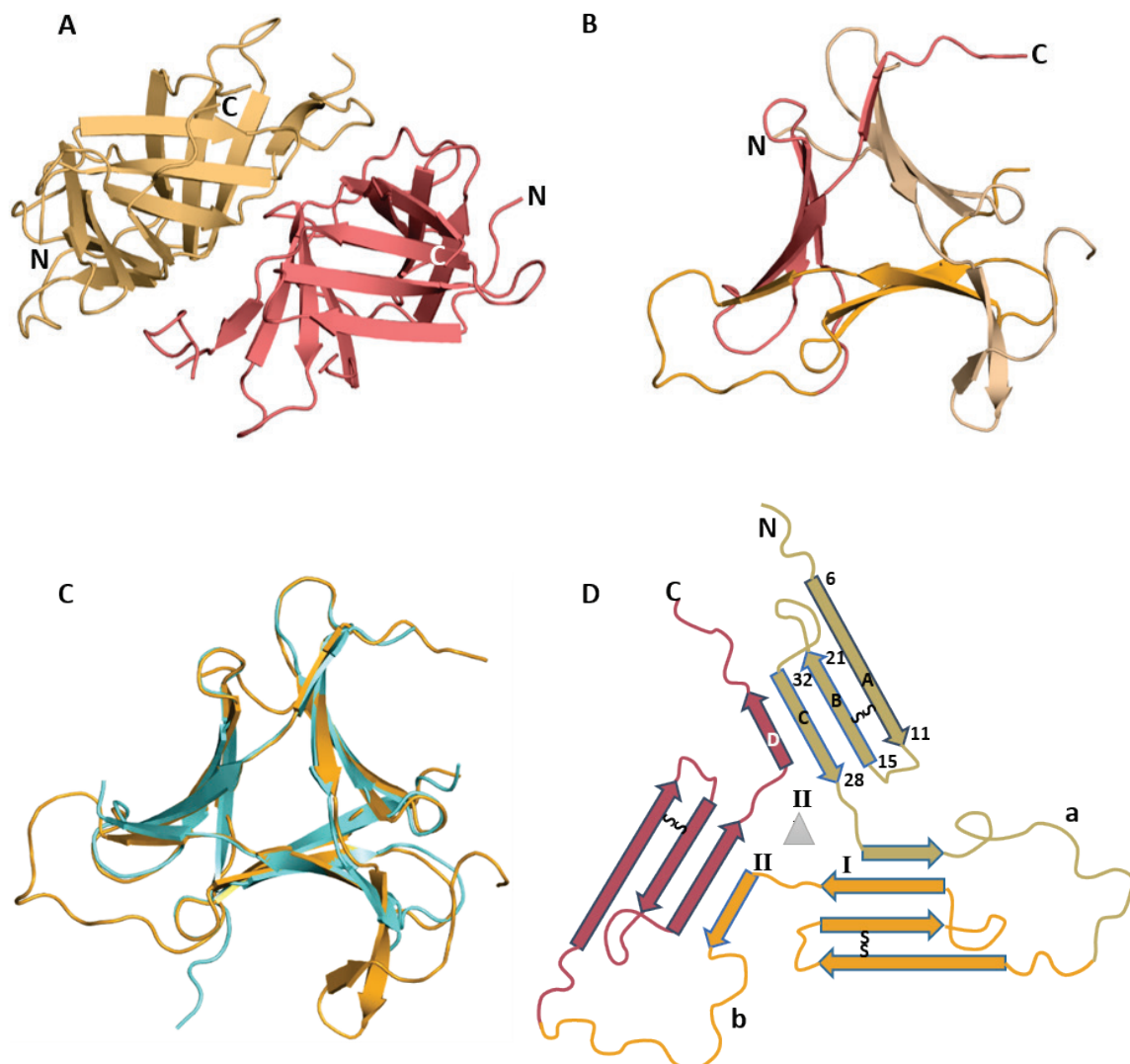


Figure IV-42. Représentation de la structure de rPhoNLSeMet résolue en co-cristallisation avec le glycan 1 (β -prisme III).

A. Représentation cartoon du dimère présent dans l'unité asymétrique. B. Représentation cartoon du monomère de rPhoNL vu de dessus (chaîne A). Chaque couleur correspond à une répétition. C. Superposition des structures β -prisme III de rPhoNLSeMet (orange) et rPhoSL (bleu). D. Diagramme topologique du monomère de rPhoNL. I, II et III correspondent aux trois répétitions de rPhoNL, la boucle **a** est désignée comme la boucle de la répétition II et la boucle **b** celle de la répétition III.

3.1.2 Description des interactions avec le N-glycane 1

La structure de rPhoNLSeMet montre un seul N-glycane en interaction avec la protéine dans le site 1, les deux autres sites d'interaction étant masqués par les grandes boucles de surface a et b (*Figure IV-43.A*). Il est possible que ces boucles soient flexibles en solution et permettent quand même l'insertion des sucres. Il aurait été intéressant d'obtenir la structure de rPhoNL en absence de ligand et/ou dans un groupe d'espace différent pour voir si ces boucles adoptent toujours la même conformation ou si elles sont capables de mouvements, mais aucun cristal n'a pu être obtenu dans ces conditions avec les criblages réalisés à ce jour. rPhoNL

présente la même architecture du site de fixation du fucose que rPhoSL avec une conservation des résidus aliphatiques et hydrophobes (équivalents à alanine 1, valine 3, la leucine 21 et la valine 141, tryptophane 28 et phénylalanine 23, chez PhoSL). La plupart des interactions sont conservées mais plusieurs différences importantes ont été observées (**Figure IV-43**). Il n'y a pas d'interaction avec l'alanine 1 chez rPhoNL, car elle se trouve placée trop loin du sucre (pour les chaînes A et B). De plus, l'aspartate 13 chez rPhoSL correspond au glutamate 118 chez rPhoNL, empêchant l'interaction avec le groupement N-acétyl du GlcNAc 2. Le glutamate 118 semble cependant adopter deux autres conformations lui permettant d'interagir soit avec le O6 du GlcNAc2 (chaîne A), soit avec le O2 du mannose 3 (chaîne B).

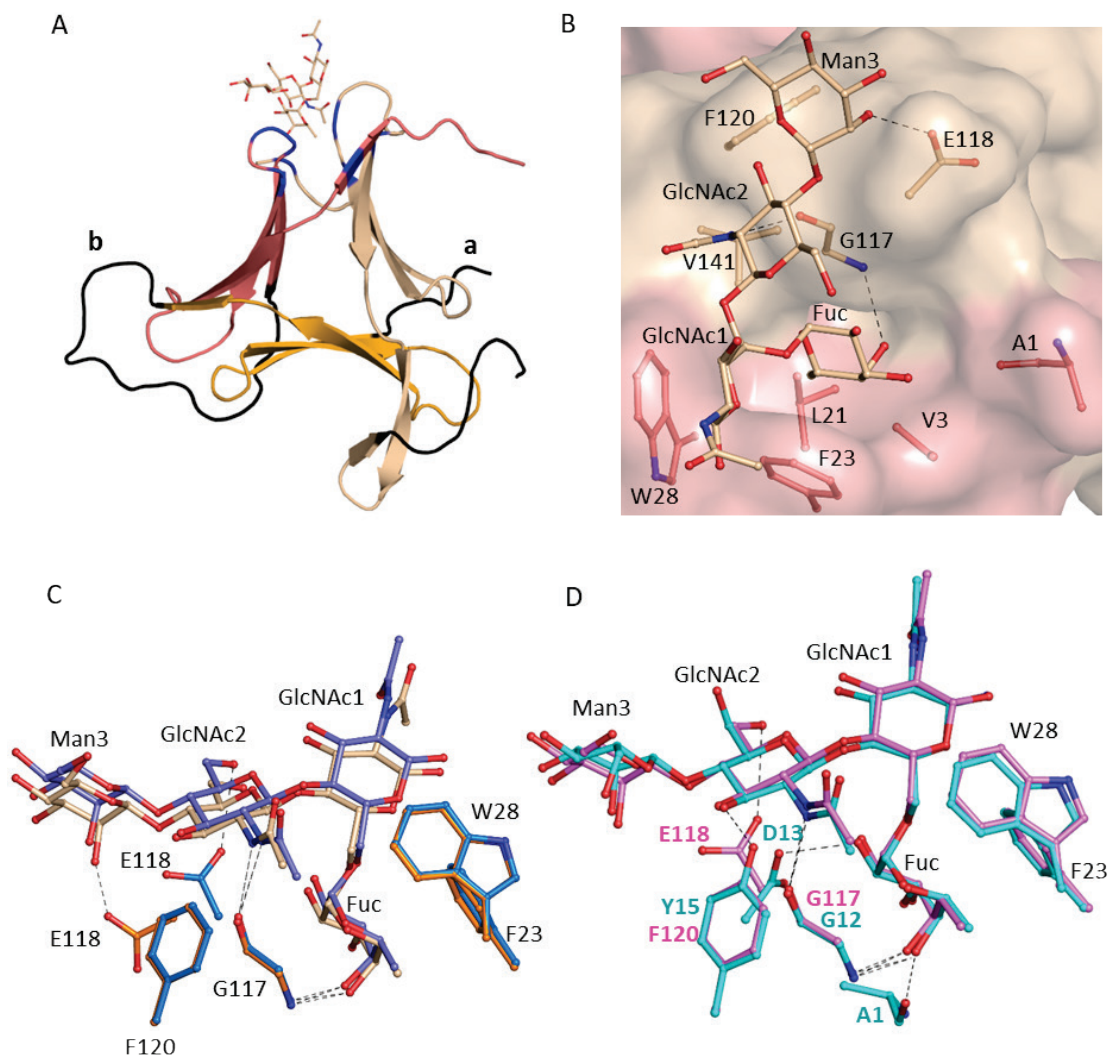


Figure IV-43. Analyse des interactions entre rPhoNLSeMet et le N-glycane 1.

A. Représentation cartoon de rPhoNL avec le N-glycane 1 en interaction avec les acides aminés indiqués en bleu. Grandes boucles de surface a et b en noir. Coloration par répétition. B. Description détaillée de l'interaction avec la poche du fucose en représentation de surface et les acides aminés impliqués dans l'interaction colorés par répétition. C. Superposition des chaînes A (bleu) et B (orange) de PhoNLSeMet. D. Superposition de la chaîne A de PhoNL (magenta) et de rPhoSL en interaction avec le N-glycane 1 (cyan).

Mais ce glutamate n'est présent que dans une seule des trois répétitions de rPhoNL, les deux autres possédant respectivement une sérine et une alanine, ce qui limite les interactions et donc la fonctionnalité de ces sites. Enfin, la tyrosine 15 qui interagissait via sa fonction hydroxyle avec le GlcNAc2 chez rPhoSL, est remplacée par une phénylalanine dans toutes les répétitions de rPhoNL, empêchant toute interaction hydrogène avec le sucre. Seulement quatre unités saccharidiques sur huit ont pu être construites en densité car la branche 1,3 exposée au solvant n'est pas du tout stabilisée par les contacts cristallins et la phénylalanine 120 stabilise probablement moins le GlcNAc5' comparée à la tyrosine 15 chez rPhoSL.

Il est probable que les interactions plus limitées chez rPhoNL suite à l'absence des résidus équivalents à l'aspartate 13 et la tyrosine 15 soient à l'origine de la moins bonne affinité de cette lectine pour les N-glycanes fucosylés comparée à rPhoSL. De plus, les deux sites d'interaction masqués par les boucles a et b sont probablement peu actifs voire inactifs pour les mêmes raisons. On peut cependant rappeler que les différences au sein du site d'interaction n'affectent pas l'architecture globale du site de fixation et donc la reconnaissance spécifique de la core-fucosylation. Seules les interactions hydrogènes sont impactées et donc l'affinité. Afin de mieux comprendre les facteurs empêchant la fixation des sucres sur les deux autres sites d'interaction, la structure de rPhoNL a été superposée à celle de rPhoSL de manière à simuler la présence des sucres sur les trois sites d'interaction de rPhoNL. Ces simulations montrent que les grandes boucles a et b interfèrent avec la fixation du N-glycane 1 et en particulier avec le fucose et le GlcNAc1 (**Figure IV-44.B**). La superposition des chaînes A et B de rPhoNL montre que les boucles ont des structures similaires, laissant supposer une relativement faible flexibilité. Néanmoins, cela peut être dû en grande partie à l'empilement cristallin, et l'on pourrait imaginer que ces boucles sont capables de bouger en solution et permettre l'insertion des sucres. Cela n'est envisageable que dans le cas où rPhoNL ne forme pas de dimères en solution. En effet, la fixation des sucres au niveau des grandes boucles a et b créerait des clashes avec le deuxième monomère. (**Figure IV-44. C et D**).

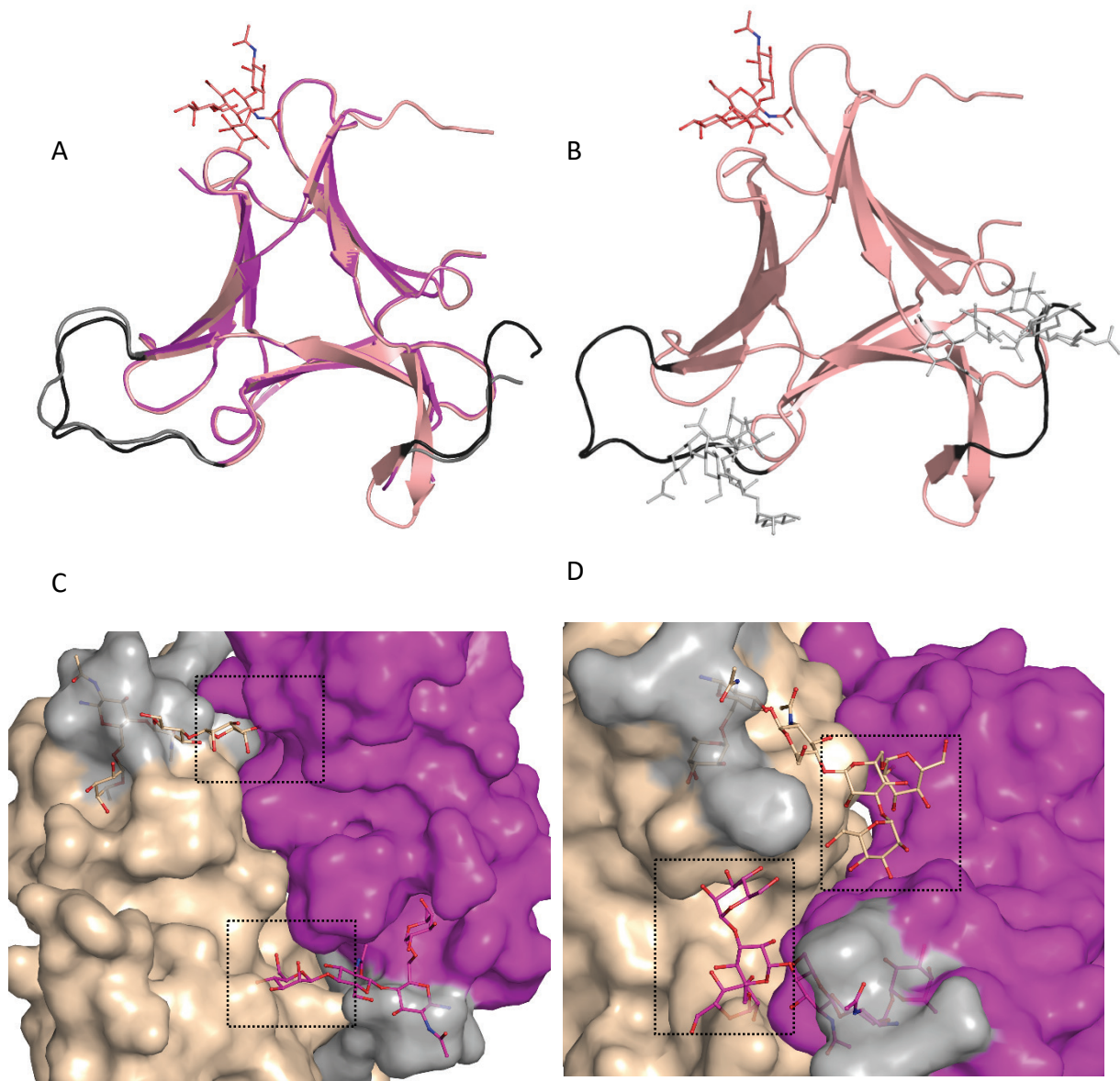


Figure IV-44. Représentation de l'encombrement stérique empêchant la fixation des sucres au niveau des boucles a et b de rPhoNLSeMet.

A. Superposition des chaînes A (beige) et B (magenta) de rPhoNLSeMet. B. Représentation cartoon d'un monomère de rPhoNL avec le N-glycane 1 en interaction. Les grandes boucles a et b sont en noir et les sucres de rPhoSL mimant l'interaction en gris C. Représentation des clashes à l'interface entre deux monomères de rPhoNL (entourés en noir). Les sucres de rPhoSL miment la position hypothétique chez rPhoNL au niveau de la grande boucle a (en vert). D. Représentation des clashes au niveau de la grande boucle b.

Le fait qu'un seul site d'interaction soit occupé chez rPhoNL pourrait donc s'expliquer à la fois par la présence des deux grandes boucles de surface, mais aussi par une tendance à la dimérisation, puisque rPhoNL forme des dimères au sein des cristaux. Cependant, cette lectine étant capable d'héماغglutiner les globules rouges de lapin, elle est forcément multivalente en solution, et est donc soit capable de dimériser sous certaines conditions, soit

capable de bouger une des grandes boucles a et/ou b pour permettre l'interaction d'au moins un deuxième sucre (ou bien les deux hypothèses simultanément).

La protéine SL2-1, strictement spécifique de la core-fucosylation (Vainauskas et al., 2016) a une séquence homologue à celle de rPhoNL et possède aussi trois répétitions reliées entre elles par deux séquences correspondant probablement à des grandes boucles comme pour rPhoNL (**Figure IV-45**). L'analyse ITC de SL2-1 a été décrite dans l'article de 2016 et a montré qu'elle avait une stoechiométrie égale à 2. Sachant que cette protéine est décrite comme formant des monomères en solution, cela signifie que deux sites d'interaction sont accessibles aux ligands, et donc qu'au moins une des deux boucles est capable de bouger pour favoriser l'insertion d'un sucre. La comparaison des séquences des boucles de rPhoNL et de SL2-1 montrent peu de similitudes. De plus, les boucles sont plus grandes pour SL2-1 (18 acides aminés pour la boucle a et 20 acides aminés pour la boucle b) que pour rPhoNL (16 acides aminés pour la boucle a et 9 acides aminés pour la boucle b).

	1	10	20	30	40	50							
PhoNL1	APVPVT	KLTC	CDGAS	SFMC	CTANL	DFADGN	WVAQW	TSPV	VFHRG	NFGAPD	HKDI	EQMAM	...
SL2-1a	APVPVT	SLNC	CDGD	TYT	CTAGL	RFGDGR	WVAQW	NVNV	VFHQ	SPASQ	QRA	ARSAD	APRGAA
PhoNL2	APVPVT	ELTC	CDGS	SFL	CTANL	DFADGN	WVAQW	PANV	VFHS	GPHPG	GGRKL	F	...
SL2-1b	APVPVT	TLTC	CDGD	TYK	CTAGL	QFGDGR	WVAQW	NANV	VFHQ	AMANA	VPTAS	AKTV	KAPREAA
PhoNL3	APVPVT	DLNC	CDGE	TFM	CTGNL	DFADGN	WKAQW	NNTN	VFHN	NNFF	FKQN
SL21-c	APVPVT	ALNC	CDGD	TYK	CTAGL	QFGDGR	WVAQW	DASV	VFHQ	S

Figure IV-45. Alignement des séquences de rPhoNL et SL2-1 par répétition (1, 2, 3 pour rPhoNL et a, b, c pour SL2-1).

Coloration par pourcentage d'identité. Séquences reliant deux répétitions consécutives entourées en noir.

Une recherche de séquence homologue complète a alors été lancée afin de déceler des acides aminés conservés au sein des boucles a et b (**Figure IV-46**). L'alignement des séquences montre des homologies entre la séquence de la boucle a de rPhoNL et les séquences correspondantes de protéines issues des organismes *Psilocybe cyanescens* et *Hypholoma sublateritium* avec la conservation des acides aminés F42, G43, D46, H47, K48 et M54. On retrouve par contre très peu d'homologies avec les séquences de protéines issues de *Streptomyces* (dont SL2-1 fait partie). Pour ce qui est de la boucle b, peu d'acides aminés sont conservés chez rPhoNL, *Hypholoma* et *Psilocybe* à part la phenylalanine 97 et la leucine 104 alors que la majorité des résidus de la boucle b de SL2-1 sont conservés chez la plupart des protéines issues de *Streptomyces*. De plus, la boucle b est plus longue chez *Streptomyces* (20 acides aminés pour les séquences les plus proches de SL2-1) que pour les autres organismes (8 à 9 acides aminés pour les séquences proches de rPhoNL). Les boucles de rPhoNL ont donc des séquences éloignées de celles de SL2-1, ce qui pourrait expliquer les différences

d'accessibilité des sucres. En effet, les différences de composition en acides aminés et de longueur des boucles peuvent avoir des conséquences sur leur flexibilité et donc leur capacité à bouger pour permettre l'insertion d'un sucre.

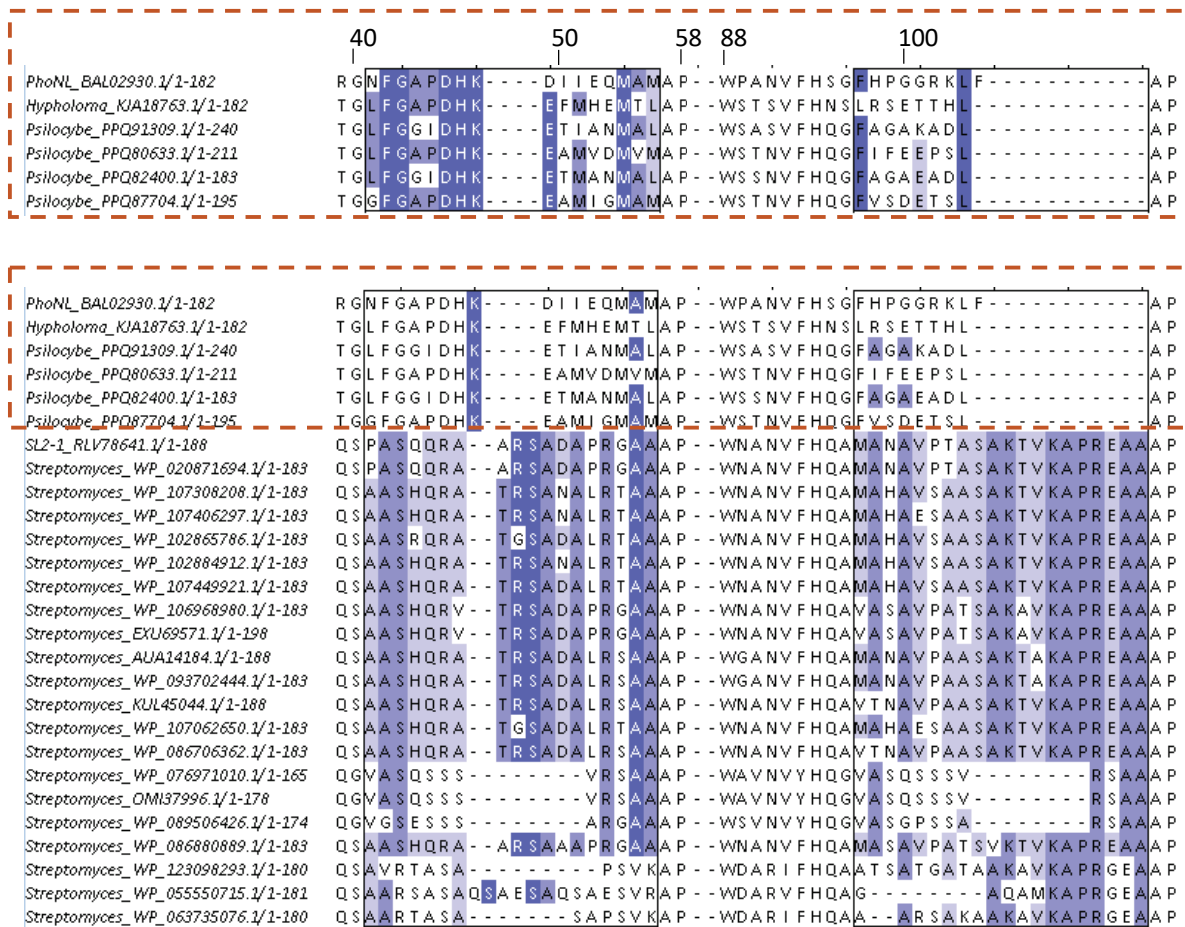


Figure IV-46. Alignement de séquences homologues à rPhoNL centré sur les grandes boucles de surface.

Les séquences ont été coupées pour centrer la figure sur les séquences des boucles a et b entourées en noir. Sont entourées en pointillés les séquences les plus proches de celles de rPhoNL. En haut, l'alignement de rPhoNL avec les séquences des boucles ayant le plus haut pourcentage d'identité. En bas, l'alignement de l'ensemble des séquences. La numérotation correspond à la séquence de rPhoNL. Alignement réalisé avec Jalview (Waterhouse et al., 2009).

3.2 Mutations pour améliorer l'affinité de rPhoNL pour ses ligands (rPhoNLm)

L'analyse des structures de rPhoNL et rPhoSL, et la comparaison avec des protéines de séquence homologue ont montré l'importance et la conservation de l'aspartate 13 et de la tyrosine 15. Ces deux acides aminés n'étant pas conservés dans la séquence de rPhoNL, cela pourrait expliquer la très faible affinité de cette lectine. Un nouveau gène a donc été synthétisé (Eurofins Genomics) afin d'avoir un aspartate et une tyrosine en positions 13 et 15 sur les trois répétitions (**Figure IV-47**).

```

PhoNL1      APVPVTKLTCDGASSEMCTANLDFADGNWVAQWTSFVFHFRGNFGAPDHKDIIEQMAM
PhoNL2      APVPVTELTCDGSSSEELCTANLDFGDGNWVAQWPANVFHSGFHPGGRKLF-----
PhoNL3      APVPVTDLNCDDGETEMCTGNLDFADGNWKAQWNTNVFHNNNFVKQN-----
PhoSL       APVPVTKLVCDGDTYKCTAYLDFGDGRWVAQWDTNVFHTG-----
***** * *** :: ** * * . * * * * * * * * *

```

Figure IV-47. Alignement des séquences de rPhoNL et rPhoSL.

En vert, les acides aminés ciblés pour la mutation de rPhoNL : A, S et E mutés en D et les trois F mutées en Y.

De plus, l'analyse de rPhoSL a permis de mettre en évidence que les trois acides aminés en N-terminal « GAM », qui ne sont pas présents dans la séquence d'origine, et sont conservés après coupure de l'étiquette à la protéase TEV, induisent une diminution de l'affinité. La lectine rPhoNL ayant été clonée dans le même vecteur, et produite de façon identique à rPhoSL, possède aussi ces trois acides aminés résiduels en N-terminal. Le gène de rPhoNL nouvellement synthétisé a donc subi également une étape de mutagenèse dirigée afin de supprimer l'alanine et la méthionine N-terminale.

3.2.1 Production et analyse biochimique de rPhoNLm

La lectine ainsi mutée (rPhoNLm) a ensuite été produite et purifiée selon les mêmes protocoles que précédemment, et a été obtenue avec des rendements similaires. L'analyse de la stabilité thermique par TSA a montré que rPhoNLm est plus stable qu'avant mutation. En effet, aucune dénaturation n'a été observée entre 0 et 100°C alors qu'auparavant rPhoNL avait un T_m autour de 70°C et présentait une dénaturation en deux étapes. La lectine présente une stabilité similaire à rPhoSL après mutation.

rPhoNLm est toujours capable d'héماغglutiner les globules rouges de lapin. Lorsque les globules rouges sont préalablement traités à la trypsine, sa capacité d'héماغglutination est 125 fois meilleure qu'avant mutation. Lorsque les globules rouges ne sont pas traités, les capacités d'héماغglutination avant et après mutation sont équivalentes (*Tableau IV-7*). Ces résultats montrent que les mutations réalisées sur rPhoNL améliorent la reconnaissance des ligands peu accessibles en surface des globules rouges mais qu'en absence de traitement, la reconnaissance reste environ 36 fois plus faible que pour rPhoSL. Il est possible que cela soit dû aux différences structurales entre rPhoSL et rPhoNL : les grandes boucles de surface de rPhoNL (et sa dimérisation potentielle) ne lui permettent peut-être pas d'accéder aussi facilement aux ligands des globules rouges que rPhoSL. Le traitement à la trypsine qui permet de dégager l'accès aux sucres a donc un plus fort impact sur la capacité d'héماغglutination de rPhoNL.

	Concentration limite d'hémagglutination (mg.mL ⁻¹)			
	rPhoNL	rPhoNLm	rPhoSL	rPhoSLm
Traitement trypsine	1	0.008	0.002	0.0009
Sans traitement	0.25	0.25	0.007	0.007

Tableau IV-7. Comparaison des capacités d'hémagglutination des globules rouges de lapin entre rPhoNL, rPhoNLm, rPhoSL et rPhoSLm.

Ensuite, des analyses par ITC ont été réalisées avec rPhoNLm pour déterminer si les mutations réalisées permettaient effectivement d'augmenter l'affinité de la lectine comme supposé. Les mesures ont à nouveau été réalisées avec le chitobiose core-fucosylé et du N-glycane 1 et ont permis cette fois de mesurer des affinités de 23 μ M et 13 μ M respectivement (**Tableau IV-8**). Ces résultats démontrent que l'aspartate et la tyrosine en positions 13 et 15 respectivement ont un impact important sur l'affinité et que les mutations réalisées pour conserver ces deux acides aminés dans les trois répétitions de rPhoNL, en plus de la mutation du N-terminal, ont suffi à obtenir une affinité équivalente à celle de rPhoSLm. On retrouve, comme précédemment avec rPhoSL, une affinité environ 1.8 fois meilleure avec le N-glycane 1 comparé au chitobiose core-fucosylé (en raison des interactions supplémentaires avec la branche α 1,6 du N-glycane 1). La stœchiométrie, de rPhoNLm n'a pas pu être déterminée au cours de ces expériences car les mesures ont dû être réalisées avec un excès de sucres, et ce mode d'analyse ne permet pas d'avoir des valeurs fiables sur la stœchiométrie. Les analyses ITC n'ont donc pas permis de savoir si plusieurs sucres étaient capables de se fixer par monomère, et donc si les grandes boucles de surface pouvaient bouger en solution.



	Affinités (μ M)	
	Chitobiose core-fucosylé 	N-glycane 1 
rPhoNL	ND (trop faible)	ND (trop faible)
rPhoNLm	23	13
rPhoSLm	10	

Tableau IV-8. Comparaison des affinités mesurées par ITC entre rPhoNL, rPhoNLm et rPhoSLm. ND = non déterminé.

3.2.2 Analyse structurale de rPhoNLm avec le N-glycane 6

Les analyses précédentes ayant démontré que les mutations réalisées sur rPhoNLm permettaient effectivement d'améliorer l'affinité de la lectine comme escompté, l'objectif a été ensuite de résoudre la structure de la lectine mutée, afin de vérifier que la mutation n'avait pas d'impact sur la structure globale et de vérifier que de nouvelles interactions se font bien

avec l'aspartate en position 13 et la tyrosine en position 15. Cela permettait également de voir si le fait d'avoir une meilleure affinité permettait aux sucres d'interagir sur les sites où les grandes boucles de surface sont présentes. La cristallogénèse de rPhoNLm a été réalisée en co-cristallisation avec le N-glycane 6, qui est un N-glycane bissecté. L'objectif était à la fois de résoudre la structure du mutant mais aussi de déterminer si des interactions supplémentaires existent avec ce type de sucres car aucun cristal n'avait encore pu être obtenu avec ce N-glycane. Des cristaux de rPhoNLm ont pu être obtenus, toujours sous forme de plaque dans des conditions similaires à celles de la lectine non mutée avec une solution contenant 12.5 % PEG 2K, 200 mM MgCl₂, 0.1 M Tris pH 7. Les données ont été collectées sur la ligne Proxima-1 à SOLEIL, Saint Aubin sur un détecteur Pilatus 6M (Dectris Ltd). La structure a été résolue dans le même groupe d'espace P2₁ que la protéine non mutée. Les détails des données de collecte et d'affinement sont indiqués dans le **Tableau IV-9**.

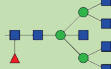
		rPhoNLm-glycane 6 	
Données collectées			
Longueur d'onde (Å) / Groupe d'espace	0.979 / P2 ₁		
Dimensions de maille a b c (Å) α β γ (°)	46.78 73.66 61.82 90 102.96 90		
Résolution (Å)	45.58-2.3 (2.38-2.3)		
Nb Réflexions mesurées /uniques	59131 / 18117		
Rmerge	0.042 (0.268)		
Mean I / sd(I)	14.2 (3.3)		
Complétude (%)	99.2 (98.9)		
Multiplicité	3.3 (3.1)		
CC1/2	0.998 (0.916)		
Affinement			
Résolution (Å)	45.58-2.3		
work/ free Reflections	18266 / 918		
R _{work} / R _{free} (%)	21.4 / 27.2		
Rmsd bond lengths (Å) / angles (°)	0.0149 / 1.81		
Nb. of atoms	Chain A	Chain B	
Protéine	1102	1057	
Ligand	88	77	
Eau	102		
B-facteurs moyens Protéine	46.9	43.8	
B-facteurs moyens Ligand	69.8	67.2	
B-facteurs moyens Eau	42.7		
Ramachandran plot (%)Allowed /Favored/ Outliers	98.9 /92.0 /3		

Tableau IV-9. Données de collecte et d'affinement pour la structure de rPhoNLm.

Les valeurs entre parenthèses correspondent à la couronne de plus haute résolution. Rfree a été calculé avec 5% des réflexions.

On retrouve comme pour rPhoNL, deux molécules dans l'unité asymétrique avec les mêmes interactions au niveau de l'interface. La superposition des structures de rPhoNL mutée et non mutée montre que les mutations réalisées ne modifient pas le repliement global de la protéine (**Figure IV-48. A**). Seuls 7 unités saccharidiques sur 11 du N-glycane 6 ont pu être construites dans la densité électronique pour la chaîne A et 6 sur 11 pour la chaîne B (Annexe **Figure VII-3**). Les sucres qui n'ont pas été construits n'interagissent pas avec rPhoNL et sont donc probablement très flexibles car exposés au solvant. Comme cela était déjà le cas pour le N-glycane 5, le mode d'interaction du N-glycane 6 est comparable à celui du N-glycane 1.

L'étude de l'interaction du N-glycane 6 avec la protéine montre qu'après mutation, rPhoNLm interagit avec l'aspartate 118 et la tyrosine 120 (correspondant aux positions 13 et 15 de la troisième répétition) de la même manière que rPhoSL. Cela confirme que l'augmentation de l'affinité après mutation est bien liée aux interactions supplémentaires avec ces deux acides aminés (**Figure IV-48. B**). On remarque également que dans la chaîne A, le fucose interagit avec la glycine 0 via son O3. Cette glycine n'a pas pu être construite en densité pour la chaîne B mais pourrait réaliser le même type d'interaction sur les deux chaînes. Il s'agit la glycine résiduelle, qui a été conservée pour permettre la coupure de l'étiquette par la protéase TEV, et cette interaction semble donc indiquer qu'elle ne gêne pas l'interaction et, au contraire, la stabilise.

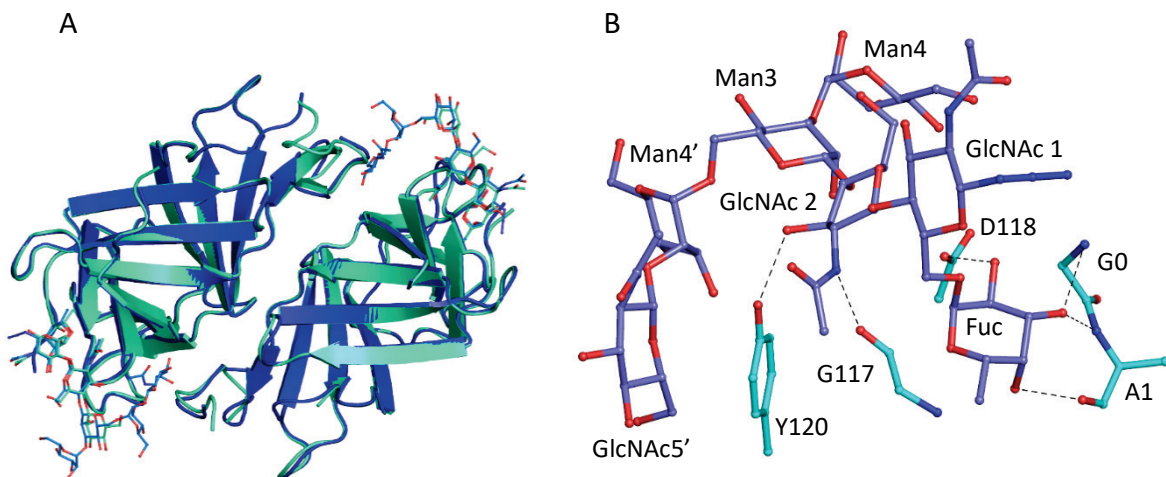


Figure IV-48. Superposition des structures de rPhoNL et rPhoNLm (A) et descriptions des interactions de la chaîne A de rPhoNLm avec le N-glycane 6 (B).

3.2.3 Analyse de l'état d'oligomérisation de rPhoNLm

Comme cela a été évoqué précédemment, l'état d'oligomérisation de rPhoNL n'a pas pu être déterminé avec certitude en raison du manque de reproductibilité des résultats qui ont

laissé supposer que rPhoNL pouvait à la fois former des monomères et des dimères en solution.

Dans le but d'obtenir des informations complémentaires sur la dimérisation potentielle de rPhoNLm, l'interface observée dans la structure cristallographique a été analysée par PDBePISA (Krissinel and Henrick, 2007). Ce programme rend compte des résidus de l'interface impliqués dans des interactions hydrogène, des ponts disulfures, et des ponts salins mais aussi des zones exposées ou non au solvant ainsi que des énergies de solvation des résidus de l'interface. Un score de 0 à 1 indique la probabilité que l'interface soit une interface de dimérisation, et ce score est de 0 pour rPhoNLm. Cela confirme que la dimérisation de rPhoNLm, si elle existe, n'est probablement pas stable. Le programme décrit cependant des interactions existant à l'interface, constituées majoritairement d'interactions hydrophobes et de deux interactions hydrogènes entre l'histidine 47 d'une chaîne et l'asparagine 138 de l'autre chaîne (*Figure IV-49*). Ces interactions ne sont probablement pas suffisantes pour stabiliser un dimère en solution.

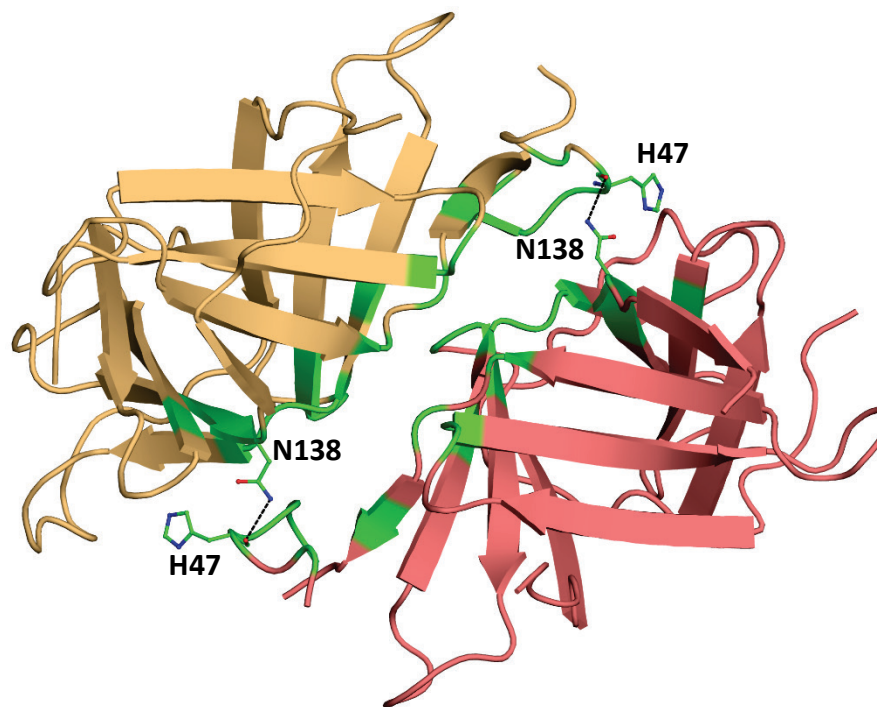


Figure IV-49. Représentation de l'interface entre les monomères de rPhoNLm.

Les résidus impliqués dans l'interaction décrits par PISA sont indiqués en vert. En noir pointillé, sont représentées les interactions hydrogènes entre les résidus H47 et N138.

Pour confirmer ce résultat dans le cas où rPhoNLm est en solution, une analyse par filtration sur gel a à nouveau été réalisée dans les mêmes conditions que ce qui avait été précédemment réalisé avec la protéine non mutée en utilisant la même colonne de filtration

sur gel (Enrich 70, Bio-Rad) (**Figure IV-50**). Selon la courbe étalon, les volumes d'éluion pour les trois pics principaux correspondent à des tailles de 26 kDa, 54 kDa et 116 kDa (pour les pics à 11.09 mL, 9.61 mL et 8.05 mL respectivement). Le pic de 116 kDa correspond vraisemblablement à des agrégats, tandis que les deux autres de 26 et 54 kDa pourraient éventuellement correspondre aux dimères et tétramères de rPhoNLm mais avec un facteur d'erreur important de 7.4 et 12.8 kDa (le monomère est de 16.7 kDa).

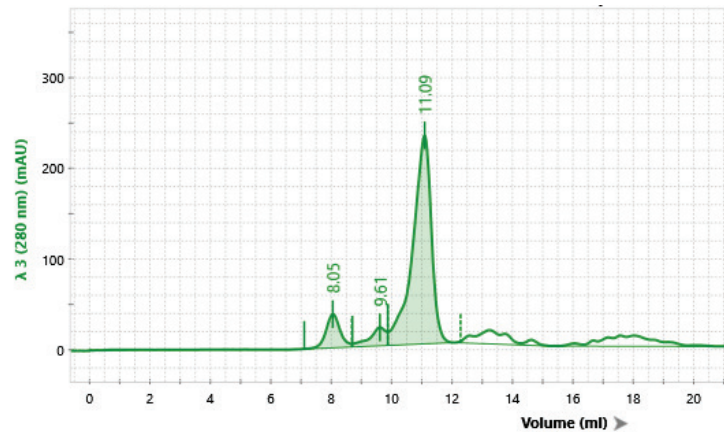


Figure IV-50. Analyse par filtration sur gel de rPhoNLm sur colonne Enrich 70. Utilisation de la colonne Enrich 70 (Bio-Rad) préalablement calibrée avec des protéines de poids moléculaire connus. Tampon 25 mM MES pH 6.

L'expérience a donc été réalisée à nouveau avec une autre colonne de filtration sur gel (S200, GE Healthcare). Cette fois-ci le résultat est complètement différent puisqu'on obtient un seul pic avec un volume d'éluion de 2.27 mL correspondant à une taille de 13.8 kDa (**Figure IV-51**). Cela semble indiquer cette fois que rPhoNLm serait monomérique avec un facteur d'erreur de seulement 2.9 kDa.

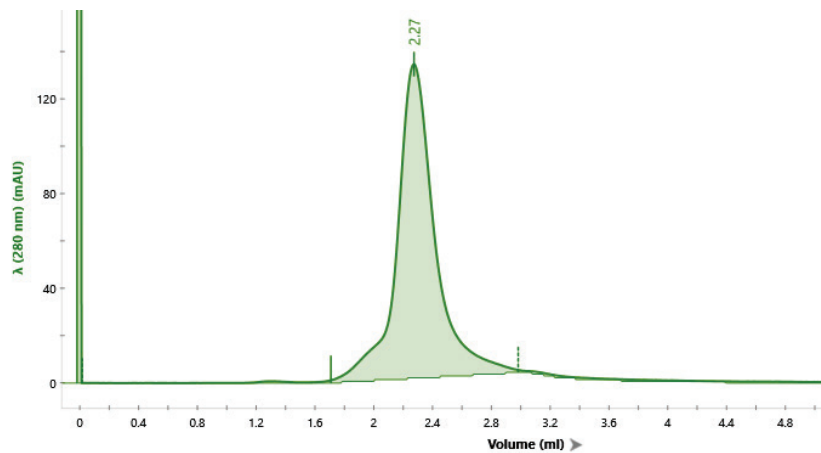


Figure IV-51. Analyse par filtration sur gel de rPhoNLm sur colonne S200. Utilisation de la colonne S200 (GE-Healthcare) préalablement calibrée avec des protéines de poids moléculaires connus. Tampon 25 mM MES pH 6.

Afin de confirmer ce deuxième résultat, et de s'affranchir des éventuels problèmes d'interactions non spécifiques sur la colonne pouvant fausser l'interprétation, rPhoNLm a finalement été testée en SEC-MALS avec une colonne S200 increase (GE-Healthcare). Le traitement des résultats indique un seul poids moléculaire de 17.1 kDa +/- 1.7%, ce qui correspond à un échantillon de rPhoNLm monomérique sans présence de contaminant ni d'agrégat (**Figure IV-52**). Ces résultats confirment que rPhoNL est monomérique en solution à la concentration testée de 3.1 mg.mL⁻¹, et que les résultats obtenus en filtration sur gel, en particulier avec la Enrich 70, sont peu interprétables. Cependant, il est toujours possible que ce résultat puisse varier dans le cas de très fortes concentrations, comme observé en cristallographie, ou selon la durée de conservation de l'échantillon.

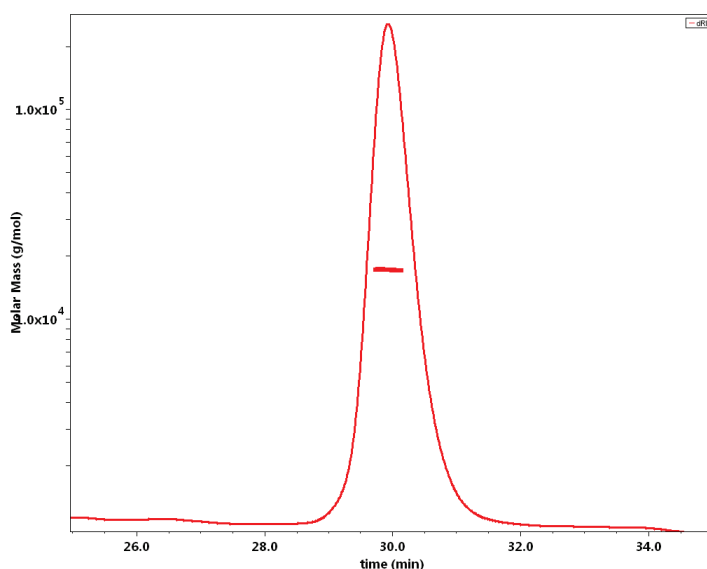


Figure IV-52. Expérience de SEC-MALS avec rPhoNLm et une colonne S200 increase. Utilisation d'un tampon 25 mM MES pH6.

3.2.4 Conclusions

La lectine rPhoNL a pu être produite sous forme recombinante dans *E. coli* avec un rendement deux fois supérieur à celui de rPhoSL. Les analyses ont montré qu'elle était très stable, et strictement spécifique de la core-fucosylation, comme attendu. Cependant, cette lectine a une affinité très faible pour les N-glycanes core-fucosylés. La comparaison de sa séquence et de sa structure avec celle de rPhoSL a permis de mettre en évidence que les sites d'interaction de cette lectine ne sont pas optimaux à cause de la non-conservation d'acides aminés fortement impliqués dans des liaisons hydrogènes. Leur composition en acides aminés qui limite le nombre d'interactions comparé à rPhoSL. La synthèse de ces informations a permis de créer un mutant, rPhoNLm, avec une affinité pour ses ligands comparable à celle de rPhoSLm. Cependant, malgré trois sites d'interaction potentiels, un seul est occupé dans la

structure cristallographique, les deux autres sites étant masqués par des grandes boucles de surface. Il est possible qu'une de ces boucles bouge en solution car l'analyse en MALS a montré que rPhoNL était monomérique et sa capacité d'hémagglutination signifie qu'elle est multivalente. L'hypothèse de la formation de dimères n'est pas totalement à exclure dans des cas particuliers de fortes concentrations comme cela a été observé en cristallographie.

3.3 Mutations pour créer un dimère stable de rPhoNL : rPhoNLmSS

La structure cristallographique de rPhoNL montre qu'il existe une interface entre deux monomères malgré le fait que les interactions soient trop faibles pour former un dimère stable. La création d'un dimère permettrait d'avoir une lectine de même spécificité que rPhoSL mais avec une organisation différente des sites d'interaction et donc potentiellement de nouvelles propriétés dans la reconnaissance de ligands sur une surface. De plus, cela permettrait d'avoir un échantillon toujours homogène en s'affranchissant d'éventuels échanges monomères-dimères en fonction de la concentration.

La stratégie adoptée pour former un dimère stable a été de rechercher les acides aminés de l'interface susceptibles de former des ponts disulfures après mutation en cystéine. Cette recherche a été réalisée avec le programme Disulfide by Design 2 (Dombkowski, 2003) et rPhoNLm. Ce programme recherche à partir de la structure au format pdb, les paires d'acides aminés à muter en cystéine afin de former des ponts disulfures (inter ou intra-chaîne). La recherche s'appuie sur les caractéristiques de la géométrie des ponts disulfures et fournit un modèle de la structure avec les ponts disulfures formés. Les résultats ont indiqué que la sérine 70 mutée en cystéine pouvait conduire à la formation d'un pont disulfure entre deux monomères. Cette mutation a été réalisée par mutagenèse dirigée.

Le mutant rPhoNLmSS a ensuite été transformé dans les cellules BL21 Star(DE3), et produit et purifié selon la même méthode que pour rPhoNLm. Cependant les rendements sont très faibles pour la fraction périplasmique de rPhoNLmSS, avec moins de 0.2 mg par litre de culture. Des tests sur gel SDS PAGE en présence ou en absence de DTT (un agent réducteur des ponts disulfures) ont été réalisés pour évaluer la formation des ponts disulfures de rPhoNLmSS. Les résultats présentés dans la **Figure IV-53** montrent qu'en absence de DTT, une deuxième bande apparaît vers 30 kDa, la taille attendue pour le dimère de rPhoNLmSS. Cette deuxième bande n'apparaît pas pour la protéine non mutée, et semble donc indiquer qu'une partie de rPhoNLmSS forme bien des dimères et qu'un pont disulfure se crée bien

entre les cystéines 70. Cependant, d'après ces résultats, les ponts disulfures ne se forment pas correctement pour plus de la moitié des protéines.

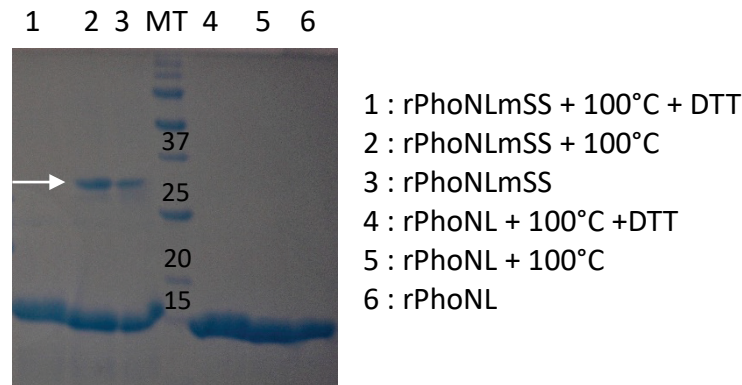


Figure IV-53. Analyse sur gel SDS-PAGE de la dimérisation de rPhoNLmSS purifiée.

MT = marqueur de taille (en kDa). 100°C indique que l'échantillon a été préalablement chauffé à cette température pendant 5 minutes. La flèche blanche indique la taille attendue pour le dimère de rPhoNLmSS.

Dans le but d'améliorer la proportion de ponts disulfures dans l'échantillon, des tests d'oxydation in vitro ont été réalisés, soit en ajoutant du glutathion oxydé et réduit (1mM GSSG + 2mM GSH) suivi d'une incubation de 72 h à température ambiante, soit en laissant incuber rPhoNLmSS seule pendant 72 h à température ambiante. Une augmentation de la proportion de ponts disulfures après 72 h d'incubation a bien été observée mais avec peu d'effet du glutathion. De plus, cette augmentation est faible et la proportion de ponts disulfure reste inférieure à 50% de rPhoNLmSS.

Étant donné les faibles rendements obtenus dans le périplasma et les difficultés de formation des ponts disulfures, la stratégie a été modifiée afin de produire rPhoNLmSS dans le cytoplasme, plutôt que dans le périplasma. Pour cela, les cellules Shuffle ont été utilisées, car il s'agit d'une souche modifiée dont l'expression de réductase dans le cytoplasme a été inhibée pour y favoriser la formation des ponts disulfures. Malheureusement, cette méthode n'a pas permis d'obtenir les résultats souhaités. Bien que rPhoNLmSS soit produite en plus grande quantité dans le cytoplasme que dans le périplasma, elle n'a pas pu être correctement purifiée sur colonne d'affinité au nickel en raison du grand nombre de contaminants présents dans le cytoplasme. De plus, la proportion de ponts disulfure correctement formés n'est pas meilleure que pour la fraction périplasmique des cellules BL21 Star toujours selon une analyse SDS-PAGE en absence d'agent réducteur.

Ces résultats rendent compte de deux problématiques qui n'ont pas été résolues : le mauvais rendement de production de rPhoNLmSS dans le périplasma et le fait qu'il y ait

moins de 50% de formation de ponts disulfures au niveau de l'interface de dimérisation. Il est possible que le problème de formation du pont disulfure soit dû à un manque d'affinité entre les monomères qui n'ont pas assez tendance à se rapprocher pour permettre aux cystéines de se rencontrer. Pour créer un dimère stable, il faudrait créer d'autres interactions à l'interface pour favoriser l'interaction entre monomères. Pour pallier au faible rendement, d'autres vecteurs permettant d'ajouter d'autres protéines-fusion que DsbA pourraient être testés. De même, d'autres cellules pourraient être utilisées comme par exemple les BL21 pLysS, recommandées pour la production de protéines toxiques.

3.4 Étude comparative des protéines strictement spécifiques de la core-fucosylation.

L'analyse des protéines homologues à rPhoSL et rPhoNL a montré que la grande majorité de ces protéines possèdent 159 à 188 acides aminés avec trois répétitions d'une séquence d'environ 40 acides aminés comme c'est le cas de rPhoNL. rPhoSL et RSL (*Rhizopus stolonifer*), constituées d'une seule répétition semblent donc être des exceptions. Les protéines fixant le fucose sont donc typiquement des protéines constituées d'une séquence à trois répétitions. Il est probable que le monomère de PhoSL soit en réalité le fragment d'une lectine monomérique constituée de trois domaines. De plus, le séquençage du gène de rPhoSL a montré qu'il codait en réalité pour une protéine de 180 acides aminés, renforçant l'hypothèse d'une protéolyse (Yamasaki et al., 2018). L'ambiguïté réside dans le fait que la séquence de PhoSL déterminée par séquençage N-terminal et que nous avons utilisée pour ce travail, ne correspond à aucune des trois répétitions du gène de PhoSL. Cette séquence a malgré tout permis la production d'une mini-lectine recombinante qui pourrait devenir un outil performant d'analyse de la core-fucosylation.

À la suite de ce travail, d'autres structures de PhoSL obtenues par RMN (résonance magnétique nucléaire) et cristallographie ont été publiées ((Yamasaki et al., 2018); (Yamasaki et al., 2019)). Ces travaux ont été réalisés à partir de la lectine synthétisée chimiquement (qu'on appellera PhoSLc), dont la structure protéique est très similaire à celle que nous avons obtenue avec la protéine recombinante (rPhoSL). Cependant, la séquence diffère légèrement de celle que nous avons utilisée (**Figure IV-54**). L'alignement des séquences indique que 4 acides aminés diffèrent entre les deux protéines aux positions 23, 27, 35 et 40. Les résidus 27 et 40 ne sont pas impliqués dans l'interaction contrairement aux résidus 35 et 23. En effet, l'asparagine 35 chez rPhoSL interagit avec l'acide sialique du N-glycane 3 et le fait d'avoir une alanine fragiliserait probablement l'interaction dans le

deuxième sillon qui stabilise la branche 1,6 de ce sucre. À l'inverse, la mutation de la phénylalanine 23 en tyrosine pourrait théoriquement permettre de créer de nouvelles interactions hydrogènes entre le OH de la tyrosine et le O5 du GlcNAc1. Afin de maximiser les interactions avec les N-glycanes, la séquence optimale de rPhoSL devrait donc avoir une tyrosine en position 23 et une asparagine en position 35.

	1	10	20	30	40					
PhoSLc	APVPVTKLVCDGDTYKCTAYLDYGDGKWVAQWDTAVFHTT									
rPhoSL	APVPVTKLVCDGDTYKCTAYLDFGDGRWVAQWDTNVFHTG									

Figure IV-54. Alignement des séquences de rPhoSL et de PhoSLc.

Une autre différence est que tous les ponts disulfures de rPhoSL entre les brins A et B sont toujours bien formés dans toutes ses structures, ce qui n'est pas le cas dans toutes les chaînes des structures de PhoSLc. Cela est probablement dû à l'empilement cristallin car la séquence est identique entre rPhoSL et PhoSLc dans cette région. En superposant les structures de rPhoSL avec celle de PhoSLc, on peut remarquer que lorsque les ponts disulfures de PhoSLc sont mal formés, la boucle qui relie les brins A et B se superpose mal avec celle de rPhoSL (*Figure IV-55*). Il semble donc que la conformation de cette boucle ait un impact important dans la formation du pont disulfure entre les brins A et B de chaque monomère.

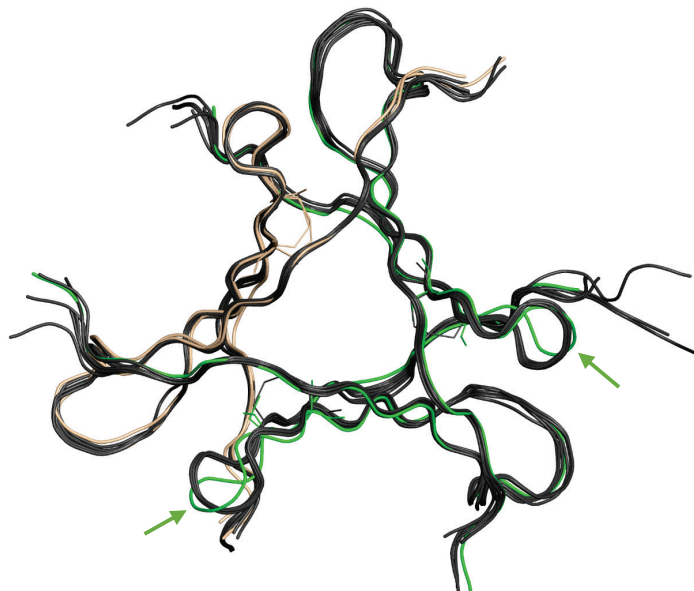


Figure IV-55. Superposition des structures de rPhoSL et de la structure de PhoSLc.

Structures de rPhoSL en noir : 6eke, 6fx1, 6fx2, 6fx3 ; structure de PhoSLc (6a86) en noir (pont disulfure formé) en beige (ponts disulfures partiellement formés) ou en vert (pont disulfure non formé) selon la chaîne. Les flèches indiquent les différences de conformation des boucles vertes par rapport aux autres structures.

Enfin, la dernière grande différence avec la structure de rPhoSL concerne l'orientation de l'alanine 1. En effet la plupart des alanines 1 des chaînes de rPhoSL n'interagissent pas avec les aspartates 11 et/ou 13 ni avec le O3 du fucose via leur azote comme c'est le cas dans la majorité des chaînes de PhoSLc. Seules les alanines 1 de la chaîne B des structures 6eke, 6fx2 et 6fx3 interagissent avec l'aspartate 11 via leur azote et seule l'alanine 1 de la chaîne I de 6fx1 interagit avec le O3 du fucose. Pour les autres chaînes, l'absence d'interaction avec les aspartates est parfois due au fait qu'elles sont trop éloignées (chaînes I, J et L de 6fx1), ou alors que l'azote est orienté dans la mauvaise direction (toutes les autres chaînes des structures de rPhoSL : 6eke, 6fx1, 6fx2, 6fx3 ; **Figure IV-56**). La différence majeure avec PhoSLc est que dans rPhoSL, l'alanine 1 est précédée des acides aminés GAM qui semblent imposer une conformation défavorable à l'interaction avec les aspartates 11 et 13. Concernant rPhoNL, les alanines sont dans la bonne orientation avant et après mutation du N-terminal (l'alanine 1 étant précédée d'une glycine après mutation), mais ne font pas d'interaction avec les aspartates car elles sont trop éloignées. Il est également à noter que toutes les alanines de PhoSLc n'interagissent pas avec les aspartates comme pour la chaîne D de 6a86 (mauvaise orientation) ou la chaîne B de 6a86 et la chaîne A de 6a87 (trop éloignée). Il semble donc qu'il y ait une certaine flexibilité du N-ter qui favorise l'interaction de l'alanine 1 avec les aspartates 11 et 13 et le O3 du fucose lorsque l'alanine est libre et défavorise ces mêmes interactions lorsque l'alanine est liée à d'autres acides aminés en N-terminal. Cependant, comme décrit précédemment, l'interaction entre l'azote de l'alanine 1 et le O3 du fucose est remplacée par l'interaction entre la glycine 0 et ce même O3 chez rPhoNLm. La stabilisation du fucose pourrait donc être équivalente entre PhoSLc et rPhoSLm et rPhoNLm.

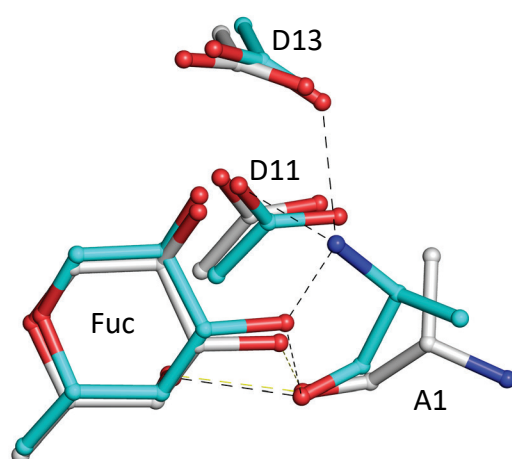


Figure IV-56. Description des orientations différentes de l'alanine 1 entre rPhoSL (6fx3 chaîne A) et PhoSLc (6a87 chaîne D).

rPhoSL est représentée en gris et PhoSLc en bleu. Les interactions sont indiquées en pointillé noir pour PhoSLc et en pointillé jaune pour rPhoSL.

Ces analyses montrent que l'affinité de rPhoSLm et rPhoNLM pourrait encore être améliorée, notamment en mutant la phénylalanine 23 en tyrosine. De plus, concernant rPhoNL, sa séquence n'est pas optimisée pour la reconnaissance de la branche 1,6 des N-glycanes dans le deuxième sillon en raison de la présence d'une proline à la place d'une asparagine en position 35 pour une des trois répétitions (l'asparagine 35 stabilise l'acide sialique chez rPhoSL). Cependant, rPhoSLm et rPhoNLM sont déjà des outils très prometteurs qui ont une stricte spécificité et une bonne affinité pour les N-glycanes core-fucosylés. Elles ont pu être produites avec succès sous forme recombinante dans *E. coli*, ce qui permettra de les utiliser facilement comme outil de marquage et de reconnaissance. De plus, cette étude a permis de définir les acides aminés impliqués dans la reconnaissance de la core-fucosylation, et les mutations de rPhoNLM ont permis d'évaluer l'impact de certains d'entre eux sur l'affinité de l'interaction.

Chapitre 4.
Étude biologique sur
cellules et tissus

4. Étude biologique sur cellules et tissus

Une fois les protéines recombinantes produites et leur spécificité pour le marqueur cancéreux d'intérêt vérifiée *in vitro* (core-fucosylation α 1,6 pour rPhoSL, rPhoSLm et rPhoNLm et antigène TF pour rTrfbL2-WT et rTrfbL2-Y111A), leur capacité de reconnaissance des motifs d'intérêt sur des cellules et tissus cancéreux a été évaluée par des expériences de FACS et de lectine-histochimie. Les détails des expérimentations sont décrits dans le *chapitre III*.

4.1 Cytométrie en flux sur lignées cancéreuses

4.1.1 PhoSL, PhoSLm et PhoNLm

Pour l'étude du marquage sur cellules cancéreuses, cinq lignées cellulaires ont été utilisées : MINO et Z138 dérivées du lymphome du manteau (MCL), HT-29 dérivée de l'adénocarcinome du côlon, MCF7 dérivées de tumeurs mammaires et A549 issues de cellules épithéliales pulmonaires adénocarcinomiques. Toutes ces lignées cellulaires sont reconnues par rPhoSLm et rPhoNLm. rPhoSL reconnaît aussi les cellules sur lesquelles elle a été testée soit MINO, Z138 et HT-29 (*Figure IV-57*).

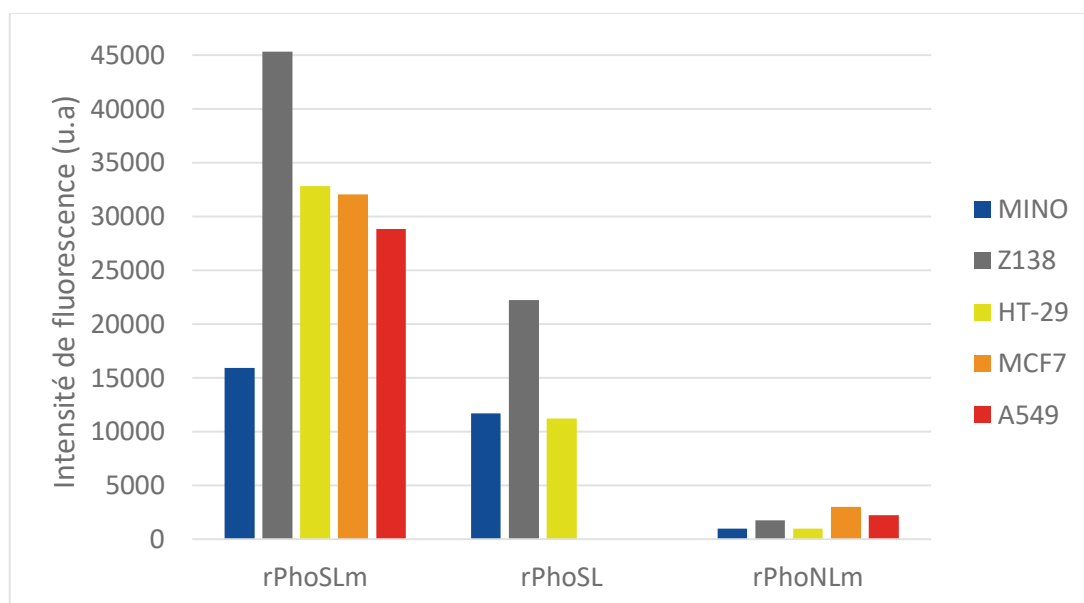


Figure IV-57. Comparaison de la reconnaissance de différentes lignées cellulaires par rPhoSL, rPhoSLm et rPhoNLm.

Les lignées MCF7 et A549 n'ont pas été testées avec rPhoSL. Lectines utilisées à $0.1 \mu\text{g.mL}^{-1}$ (MINO et Z138), $0.2 \mu\text{g.mL}^{-1}$ (HT-29) et $2.5 \mu\text{g.mL}^{-1}$ (MCF7 et A549). Les lectines ont été marquées à l'Alexa 680 (MCF7 et A549) ou avec de la biotine reconnue par la streptavidine-phycoérythrine (MINO, Z138 et HT-29).

On remarque que rPhoSLm marque davantage les cellules que rPhoSL. Ces résultats sont cohérents avec les résultats des tests d'hémagglutination qui avaient montré une capacité 2 fois meilleure pour rPhoSLm comparé à rPhoSL. La mutation du N-terminal permet donc aussi un meilleur marquage des cellules cancéreuses. rPhoNLM a un marquage beaucoup plus faible en moyenne que rPhoSL en général. Les tests d'hémagglutination avaient montré une capacité 35 fois plus faible que rPhoSL en absence de traitement avec de la trypsine et l'on retrouve cette tendance pour le marquage des lignées cellulaire cancéreuses. Cela peut être dû à la présence des boucles de surface reliant les répétitions chez rPhoNLM qui gêneraient l'accessibilité à des ligands peu accessibles comparé à rPhoSL.

Dans le but de vérifier la spécificité de la reconnaissance, c'est à dire le fait que ces lectines reconnaissent les ligands core-fucosylés sur les cellules, des premiers tests ont été réalisés en utilisant des siRNA (ON-TARGETplus Human FUT8 (2530) siRNA-SMART pool, Dharmacon) permettant d'inhiber l'expression du gène de la fucosyltransférase 8 (FUT8), l'enzyme qui catalyse le transfert de fucose lié en $\alpha 1,6$ sur les N-glycanes. Les siRNA (petits ARN double brin) ainsi transfectés dans les cellules cibles dirigent le clivage des ARN complémentaires (ceux de fut8 dans notre cas), empêchant la synthèse de l'enzyme. L'expérience a été réalisé sur les cellules MCF7 et A549. Les cellules ont été transfectées à 30-50 % de confluence avec 50 nM de siRNA FUT8 ou de siRNA LUC (luciférase, contrôle positif de la transfection) par condition, en utilisant du jetPRIME (Poluplus) comme agent de transfection. Les cellules ont été incubées avec les siRNA à 37°C pendant 72h. Les lectines ($2.5\mu\text{g}/\text{mL}^{-1}$) ont ensuite été ajoutées aux cellules suivant le protocole décrit dans le **Chapitre III**. Le contrôle avec les siRNA LUC indique que leur transfection a bien fonctionné puisqu'on a bien une extinction de la bioluminescence dans le puits contenant ces siRNA. Cependant, aucune action des siRNA FUT8 n'a été observée, les lectines reconnaissant toujours les cellules après traitement (**Figure IV-58**).

Plusieurs facteurs peuvent faire varier l'efficacité des siRNA tels que leur concentration, le choix de l'agent de transfection ou la durée d'incubation et pourraient expliquer ces résultats. Nous n'avons malheureusement pas eu l'expertise et le temps de continuer de mettre au point le protocole et poursuivre ces travaux. Afin de vérifier que les lectines rPhoSL, rPhoSLm et rPhoNLM reconnaissent bien la core-fucosylation sur les cellules et que les résultats négatifs d'inhibition avec les siRNA sont dus à un protocole inadapté, d'autres expériences d'inhibition ont été réalisées en utilisant du 6-Alkynyl-Fucose (6-AF, (Kizuka et al., 2017)) ou du chitobiose fucosylé en $\alpha 1,6$.

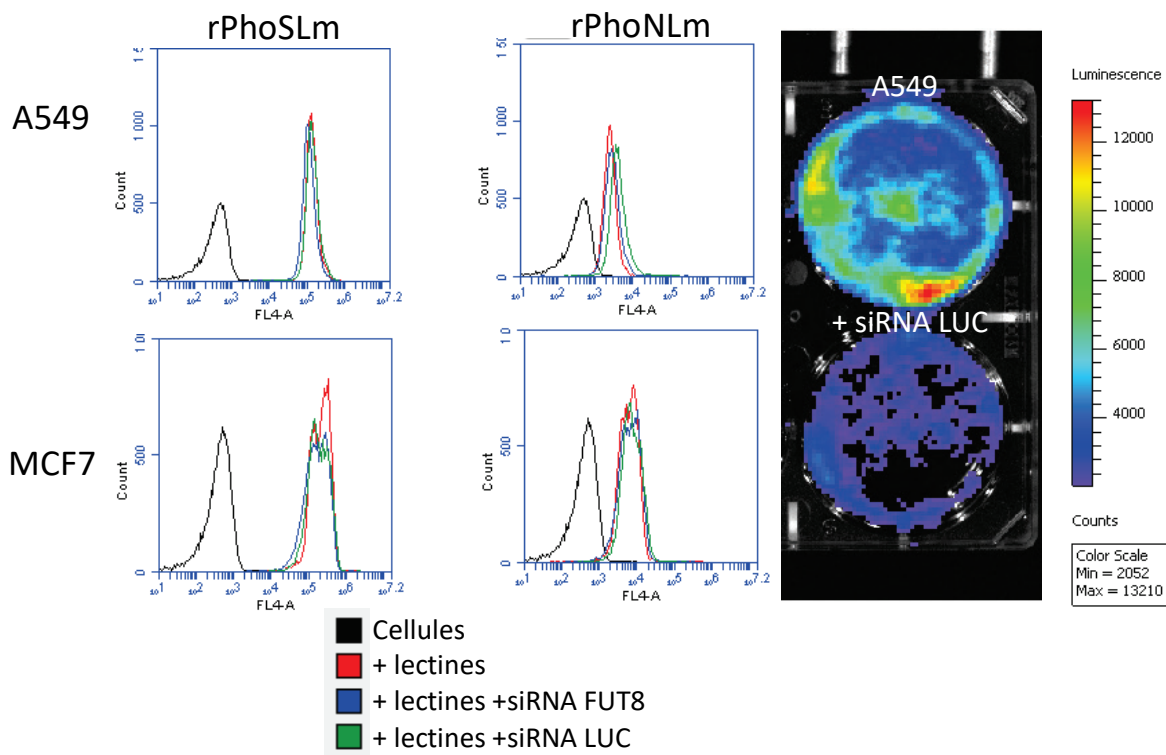


Figure IV-58. Inhibition de la fucosylation α 1,6 des N-glycanes sur les cellules A549 et MCF7 avec des siRNA FUT8.

Résultats de FACS sur cellules MCF7 et A549 avec ou sans siRNA FUT8. À droite, le résultat du contrôle de la transfection avec les siRNA LUC sur cellules A549. Lectines à $2.5 \mu\text{g.mL}^{-1}$ et marquées à l'Alexa 680.

Le 6-AF est un inhibiteur de la synthèse du GDP-fucose et a donc pour conséquence d'empêcher l'ajout de fucose sur les N-glycanes. Il agit par compétition en se fixant dans la poche catalytique de l'enzyme FX responsable de la biosynthèse du GDP-fucose. Le 6-AF a été utilisé à $20 \mu\text{M}$ et ajouté aux cellules pendant 48h avant l'ajout des lectines en renouvelant le milieu et l'inhibiteur au bout de 24h. Le chitobiose core-fucosylé a été directement mis en présence des lectines à 25 mM pendant 1 h à température ambiante et le mélange a ensuite été ajouté aux cellules. Les premières expériences ont été réalisées sur les cellules MINO, Z138 et HT-29 (*Figure IV-59*).

Les résultats d'inhibition avec une gamme de 6-AF de 0 à $20 \mu\text{M}$ montrent que l'inhibition de la fixation de rPhoSL est quasiment totale avec $2,5 \mu\text{M}$ de 6-AF (90 % d'inhibition). Pour rPhoSIm, ce niveau d'inhibition est atteint avec $5 \mu\text{M}$ de 6-AF (91 % d'inhibition). Ces résultats sont cohérents avec le fait que rPhoSIm fixe mieux les cellules que rPhoSL et semblent indiquer que ces lectines reconnaissent bien spécifiquement la core-fucosylation sur ces cellules. Concernant rPhoNLm, comme évoqué précédemment, la reconnaissance est

beaucoup plus faible que pour les deux autres lectines et l'inhibition est plus rapide avec 99 % atteints en présence de 5 μM de 6-AF, ce qui semble indiquer là aussi que PhoNLm est capable de marquer spécifiquement la core-fucosylation. Le 6-AF ne permet sûrement pas de supprimer 100 % des fucoses des N-glycanes, ce qui explique que l'on n'ait rarement 100 % d'inhibition, en particulier lorsque la reconnaissance est forte comme avec rPhoSLm. Les tests d'inhibition avec le core-chitobiose sur cellules MINO et Z138 ont permis de montrer l'efficacité de ce sucre en tant qu'inhibiteur de la fixation de rPhoSL. En effet, 25 mM de chitobiose core-fucosylé suffisent à inhiber la reconnaissance des cellules MINO et Z138 de 99 % (*Figure IV-59. B*).

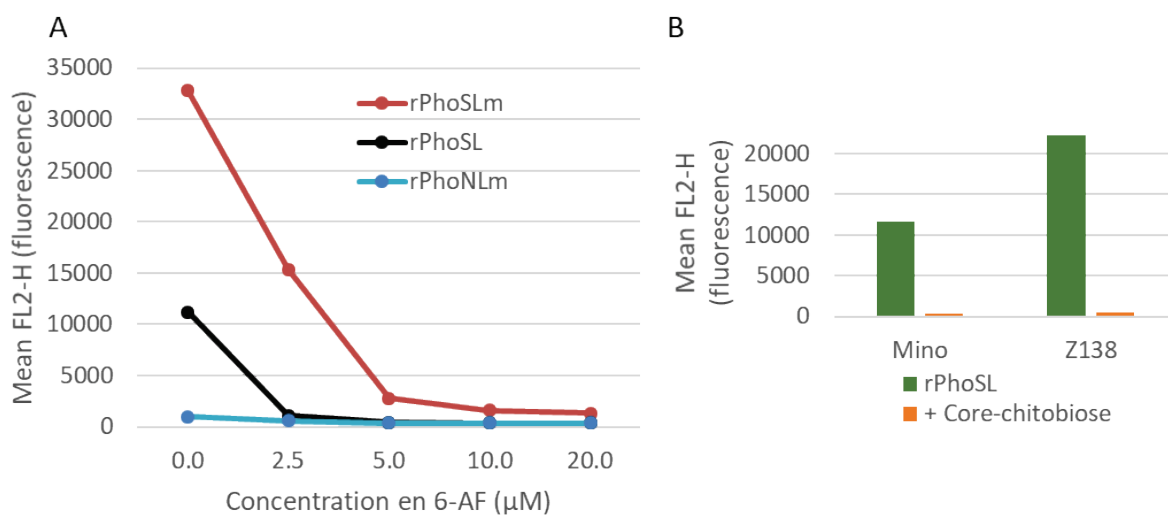


Figure IV-59. Inhibition de la fixation de rPhoSLm, rPhoSL et rPhoNLm sur cellules MINO, Z138 et HT-29 en utilisant le 6-AF ou du chitobiose core-fucosylé.

A. Inhibition de la fixation des lectines sur cellules HT-29 en utilisant une gamme de concentration de 6-AF. B. Inhibition de la reconnaissance de rPhoSL sur cellules MINO et Z138 en utilisant 25 mM de chitobiose core-fucosylé. Lectines à $0.1 \mu\text{g}\cdot\text{mL}^{-1}$ (MINO et Z138) et $0.2 \mu\text{g}\cdot\text{mL}^{-1}$ (HT-29) biotinylées puis marquées à la streptavidine-phycoérythrine.

Ces résultats étant encourageants, l'inhibition de la fixation de rPhoSLm et rPhoNLm avec le 6-AF a été réalisée sur deux autres lignées cellulaires : A549 et MCF7 en utilisant l'inhibiteur à 20 μM . Cependant, les résultats obtenus indiquent une plus faible efficacité du 6-AF sur ces lignées avec une inhibition de la fixation de rPhoSLm de 76 % sur cellules MCF7 et 30 % sur cellules A549 et une inhibition de la fixation de rPhoNLm de 64 % pour MCF7 et 34 % pour A549 (*Figure IV-60*). L'expérience réalisée en parallèle sur les cellules HT-29 prouve que l'efficacité du 6-AF n'est pas en cause puisqu'on retrouve les pourcentages d'inhibition obtenus précédemment avec cette lignée et 20 μM d'inhibiteur (98 % pour

rPhoS_Lm et 88 % pour rPhoN_Lm). Il semble donc que l'efficacité du 6-AF soit dépendante de la lignée cellulaire utilisée.

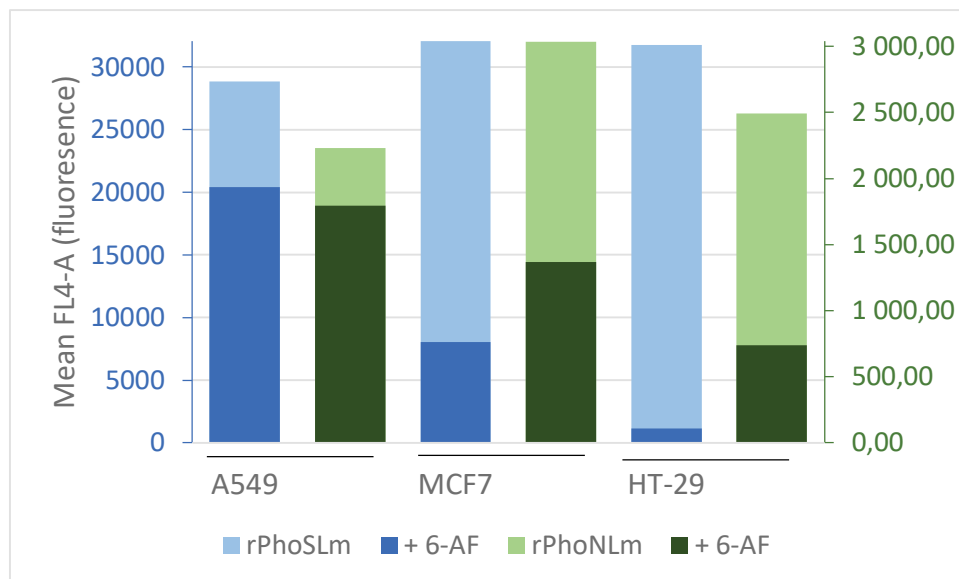


Figure IV-60. Inhibition de la fixation de rPhoS_Lm et rPhoN_Lm sur cellules MCF7, A549 et HT-29 en utilisant 20 μM de 6-AF.

Deux échelles différentes ont été utilisées pour rPhoS_Lm et rPhoN_Lm afin de mieux visualiser l'effet de l'inhibition par le 6-AF pour les deux lectines. Lectines à 0.2 μg.mL⁻¹ (rPhoS_Lm) et 2.5 μg.mL⁻¹ (rPhoN_Lm) et marquées à l'Alexa 680.

rPhoS_Lm et rPhoN_Lm sont donc capables de reconnaître les cinq lignées cellulaires testées et semblent fixer spécifiquement la core-fucosylation sur ces cellules. Comme prévu, rPhoS_Lm est moins efficace que rPhoN_Lm pour le marquage de cellules cancéreuses du fait de sa plus faible affinité pour ses ligands et le mutant est donc à privilégier pour ce type d'étude. rPhoN_Lm est beaucoup moins sensible que rPhoS_Lm, et ne semble donc pas être la lectine la plus adaptée au marquage de cellules cancéreuses.

4.1.2 rTrfbL2 et rTrfbL2-Y111A

L'analyse du marquage de cellules cancéreuses a aussi été réalisé avec les lectines rTrfbL2 et rTrfbL2-Y111A en utilisant la lignée MCF7. Cette lignée a été choisie car il a été montré qu'elle était riche en antigène TF (Rivinoja et al., 2006) et en GlcNAc terminaux (Jeong et al., 2012), les ligands reconnus par ces lectines. Les deux lectines se fixent bien sur cette lignée cellulaire avec un signal maximal atteint à partir de 5 μg.mL⁻¹ de lectine. Aucune différence significative de marquage n'est observée entre rTrfbL2 et le mutant Y111A (**Figure IV-61**). Des expériences complémentaires ont donc été réalisées afin d'évaluer la spécificité de

reconnaissance de l'antigène TF sur ces cellules, et évaluer l'impact de la reconnaissance du GlcNAc sur le site d'interaction 2.

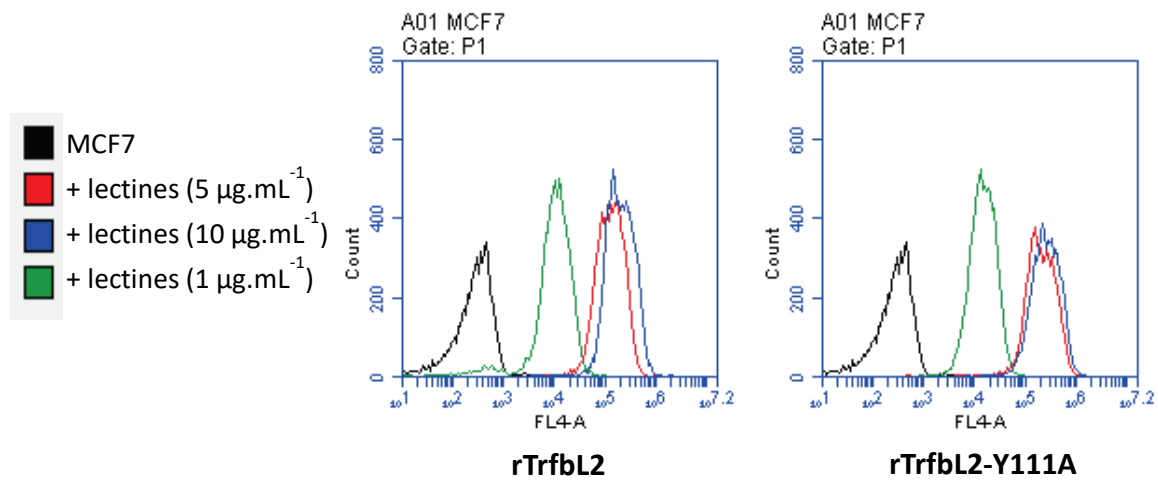


Figure IV-61. Étude de la fixation de rTrfbL2 et rTrfbL2-Y111A sur les cellules MCF7 en fonction de la concentration par la méthode FACS.

Trois concentrations de lectines ont été utilisées (1, 5 et 10 $\mu\text{g.mL}^{-1}$). Lectines marquées à l'Alexa 680.

Une forte inhibition de l'interaction est observée lorsque les protéines sont pré-incubées avec 25mM d'antigène TF (94 % et 92 % d'inhibition pour rTrfbL2 et rTrfbL2-Y111A respectivement, *Figure IV-62*). Cette inhibition est du même ordre de grandeur en présence de 25 mM d'antigène TF + 200 mM de GlcNAc (95 % pour rTrfbL2 et 92 % pour rTrfbL2-Y111A). L'ajout de 200 mM de GlcNAc seul ou de 200 mM de GlcNAc + GalNAc n'a pas d'effet significatif puisqu'on observe une faible diminution du signal pour rTrfbL2 mais une augmentation du signal pour rTrfbL2-Y111A par rapport aux mesures réalisées avec les lectines seules.

Les résultats de cytométrie en flux réalisée avec rTrfbL2 et rTrfbL2-Y111A semblent donc indiquer que ces lectines reconnaissent bien l'antigène TF sur ces cellules. L'absence d'effet significatif de l'inhibition avec du GlcNAc semble indiquer que la faible affinité de rTrfbL2 pour ce sucre sur le site 2 d'interaction n'a pas d'effet significatif sur le marquage des cellules cancéreuses. Cela peut être dû à la faible affinité pour le GlcNAc comparé à l'antigène TF, conduisant rTrfbL2 à se fixer préférentiellement sur l'antigène TF lorsque les deux sucres sont présents sur les cellules. La proportion de GlcNAc terminaux et d'antigènes TF accessibles en surface des cellules MCF7 peut aussi être en cause, et une analyse complémentaire de l'impact de l'affinité pour le GlcNAc de rTrfbL2 sur d'autres lignées cellulaires serait pertinente.

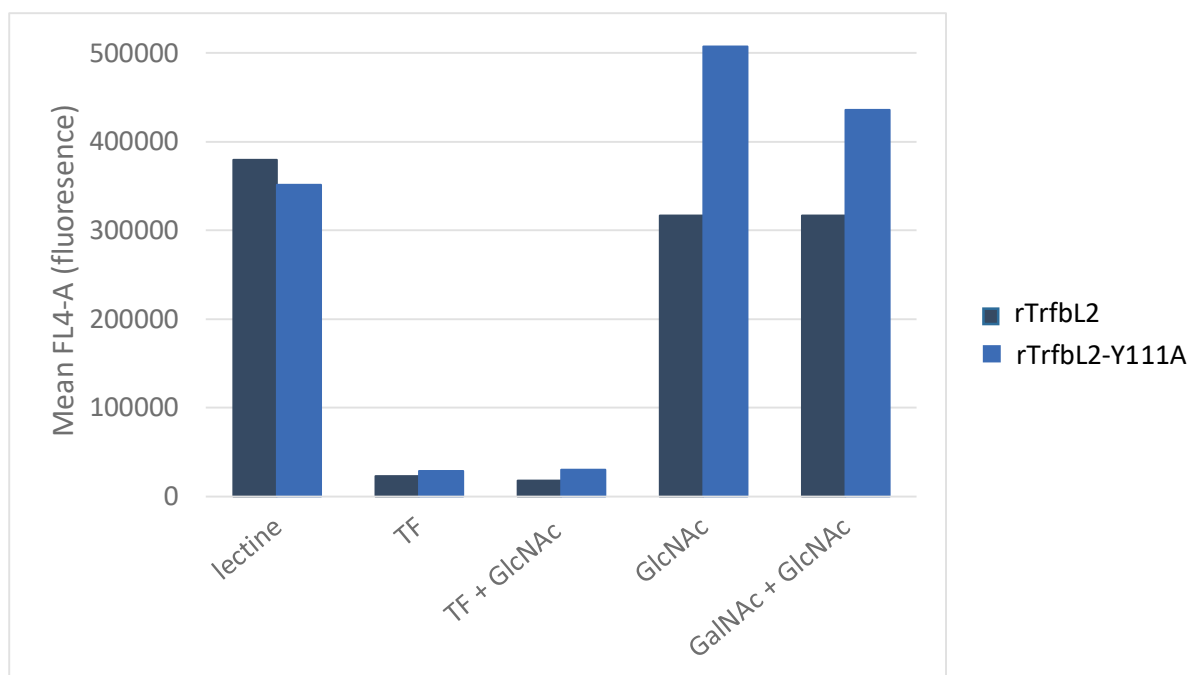


Figure IV-62. Étude de la spécificité de reconnaissance de l'antigène TF sur les cellules MCF7 pour rTrfbL2 et rTrfbL2-Y111A.

rTrfbL2 et rTrfbL2-Y111A ont été utilisées à $5\mu\text{g.mL}^{-1}$, l'antigène TF à 25 mM et le GlcNAc et le GalNAc à 200 mM. Lectines marquées à l'Alexa 680.

4.1.3 Conclusion

Les lectines rPhoSL, rPhoSLm, rPhoNLm, rTrfbL2 et rTrfbL2-Y111A reconnaissent toutes les lignées cellulaires cancéreuses testées et le marquage est dépendant de la reconnaissance spécifique de leurs ligands respectifs. La lectine rPhoSLm est la lectine la plus efficace pour la reconnaissance spécifique de la core-fucosylation sur ces cellules. Pour ce qui est de la reconnaissance de l'antigène TF, les deux lectines rTrfbL2 et rTrfbL2-Y111A présentent une efficacité similaire, et semblent se fixer préférentiellement sur l'antigène TF sans que leur faible affinité pour le GlcNAc n'apporte un biais sur le marquage.

4.2 Cytométrie en flux sur cellules mononucléées du sang périphérique

Des expériences ont été réalisées pour évaluer la capacité de rPhoSL à reconnaître les cellules mononucléées du sang périphérique (PBMC) saines. En effet, des changements du taux de core fucosylation à la surface de ces cellules peut être observé dans certaines situations pathologiques. Il a par exemple été montré que les niveaux de core-fucosylation des lymphocytes T augmentait dans le cas de colite chez la souris ou de maladie inflammatoire de l'intestin (IBD) chez l'Homme (Fujii et al., 2016). Pour cette étude, les PBMCs ont été préalablement marqués avec un mélange contenant 1/100 d'anti-CD36PE-cyanin7, 1/20 d'anti-CD19-BV605, 1/20 d'anti-CD14-FITC (BD bioscience) et avec rPhoSL-Alexa680 à 20

$\mu\text{g.mL}^{-1}$. Les résultats présentés dans la **Figure IV-63** montrent que rPhoSL se fixe sur tous les PBMCs détectés : lymphocytes T, lymphocytes NK, lymphocytes B, cellules myéloïdes, granulocytes et cellules dendritiques, indiquant la présence de core-fucosylation sur ces cellules. rPhoSL pourrait donc être aussi utilisée comme outil d'étude des variations des taux de core-fucosylation des PBMCs. De plus, la concentration de lectines utilisées pourrait être diminuée car les signaux sont déjà très élevés et il est probable que rPhoSLm soit encore plus efficace, comme observé sur les cellules cancéreuses.

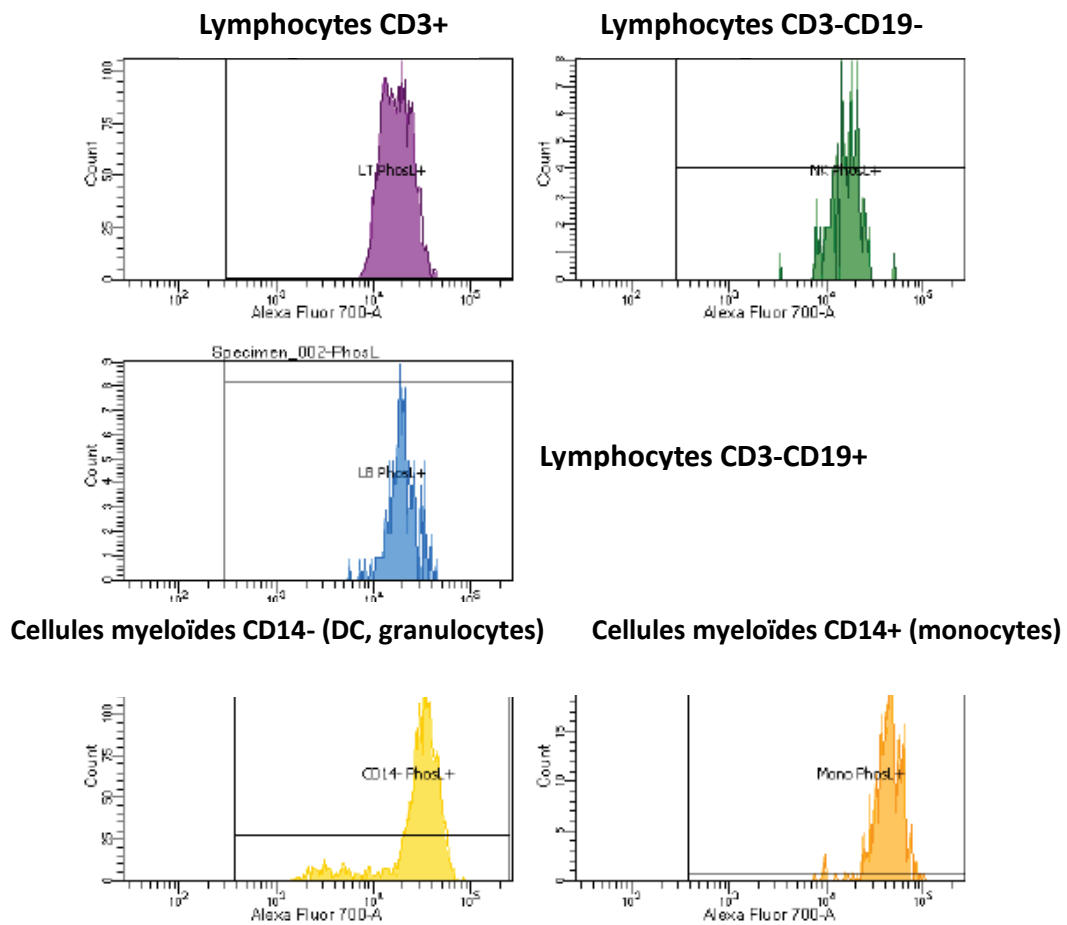


Figure IV-63. Expérience de cytométrie en flux de la reconnaissance de PBMCs par rPhoSL.

La présence ou l'absence des marqueurs CD permet de définir les sous-populations de cellules (lymphocytes T, B et NK, granulocytes, cellules dendritiques (DC), et cellules myéloïdes). rPhoSL à $20 \mu\text{g.mL}^{-1}$ marquée à l'Alexa 680.

4.3 Lectine-histochimie

Le potentiel de reconnaissance de leur ligand sur des tissus cancéreux a ensuite été évalué pour chaque lectine. Cette propriété est en effet intéressante dans l'optique de développer des outils de marquage spécifiques et polyvalents.

4.3.1 Tests sur tissus humains

Les premiers tests ont été réalisés sur rPhoSL, rPhoSLm et rPhoNLM sur des coupes paraffinées de tissus humains issus de tumeurs du côlon pour comparer les capacités de marquage de ces trois lectines par Adrien Breiman au CRCNA, Nantes. Pour ces tests, les lectines ont été préalablement biotinylées puis marquées avec de la streptavidin-péroxydase et ont été utilisées à une concentration de $50 \mu\text{g.mL}^{-1}$. Aucun marquage n'a pu être observé dans ces conditions avec rPhoSL comparé aux deux autres lectines, ce qui confirme à nouveau l'intérêt de la mutation du N-terminal de rPhoSL pour augmenter ses capacités de reconnaissance de ses ligands (**Figure IV-64**). Étonnamment, les images indiquent également un meilleur marquage avec rPhoNLM comparé à rPhoSLm, contrairement à ce qui avait été observé sur les cellules cancéreuses en solution. Cela peut être lié au fait que rPhoNLM n'a pas ses trois sites d'interaction disponibles contrairement à rPhoSLm, ce qui faciliterait la bonne orientation et la stabilisation de la lectine sur les tissus.

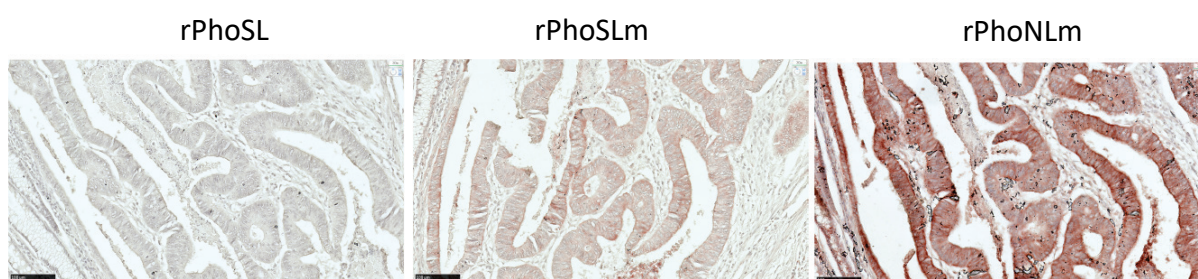


Figure IV-64. Marquage de coupes de tissus humains issus de cancers du côlon par rPhoSL, rPhoSLm et rPhoNLM.

Les lectines ont été préalablement marquées à la biotine et utilisées à $50 \mu\text{g.mL}^{-1}$ avant l'ajout de streptavidine-péroxydase.

Afin de vérifier que le meilleur marquage obtenu avec rPhoNLM n'est pas dû à des interactions non spécifiques, une analyse complémentaire a été réalisée en utilisant du chitobiose core-fucosylé en tant qu'inhibiteur. Ce test a été aussi réalisé sur des tumeurs humaines coliques mais rPhoNLM, préalablement biotinylée, a cette fois été marquée avec un fluorophore (streptavidine-Dylight 549, Vector Laboratories, Burlingame, CA). Les résultats montrent que lorsque rPhoNLM est préalablement incubée avec 25 mM de core-chitobiose, la lectine ne se fixe plus sur les tissus, ce qui semble indiquer qu'elle reconnaît spécifiquement la core-fucosylation sur ces tissus (**Figure IV-65**).

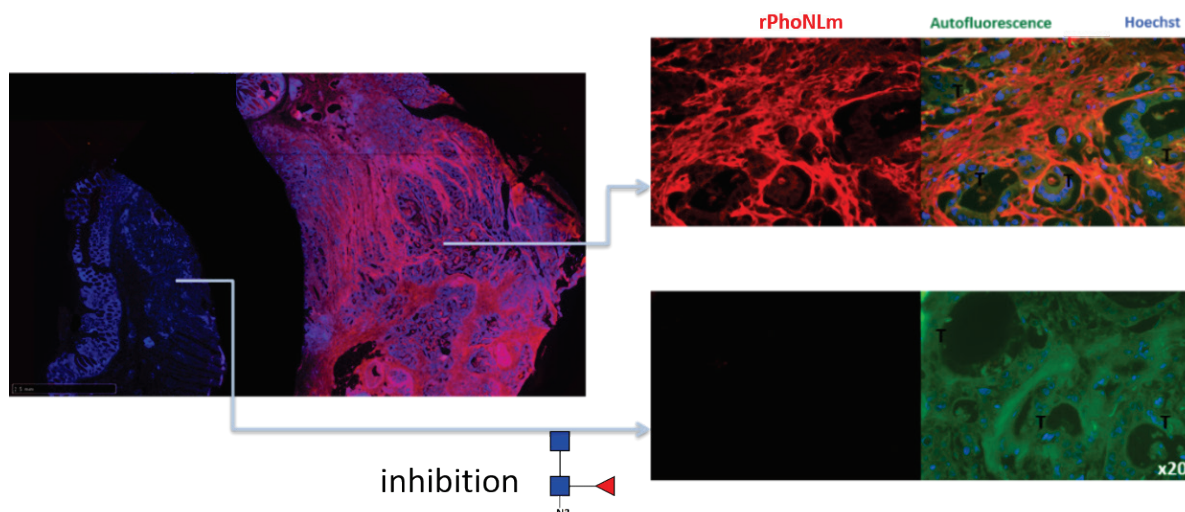


Figure IV-65. Tests d'inhibition du marquage de coupes tumorales coliques par rPhoNLm.

Inhibition avec 25 mM de chitobiose core-fucosylé. Utilisation de rPhoNLm à $20 \mu\text{g}\cdot\text{mL}^{-1}$ marquée à la biotine avant l'ajout de streptavidine conjuguée à l'Alexa 549. Le marquage de rPhoNLm est en rouge, l'autofluorescence en vert et les noyaux en bleus. T= cellules tumorales.

Des tests complémentaires ont ensuite été réalisés avec rPhoNLm sur des tissus sains dans le but de voir quels types de cellules étaient reconnus par cette lectine. Le marquage du tissu sain colique a également été comparé au tissu tumoral. rPhoNLm a été choisie pour ces tests car c'est elle qui montre le plus fort marquage. Un fort marquage du mucus au niveau du tube digestif, de l'épithélium et des glandes trachéales est observé (*Figure IV-66*). Certaines cellules spécifiques sont également marquées telles que les cellules de Kupffer dans le foie et les chondrocytes dans le cartilage de la trachée. Ces marquages montrent que la core-fucosylation est présente sur tous les tissus testés mais avec des distributions et une concentration spécifique à chaque tissu. Concernant la comparaison des tissus sains et tumoraux coliques, le marquage est très fort sur les tissus sains, et reste fort sur les tissus tumoraux, il est donc difficile de conclure sur une augmentation ou une baisse de la core-fucosylation due à la présence de tumeurs. Les différences de marquages semblent reposer essentiellement sur les différences de structures des deux tissus. Il sera cependant judicieux de diminuer les concentrations de rPhoNLm pour les études à venir car la forte intensité des marquages avec une concentration de $20 \mu\text{g}\cdot\text{mL}^{-1}$ ne permet pas d'interpréter correctement ces résultats. De plus, il serait intéressant d'être en mesure de réaliser cette comparaison sur des tissus issus d'autres organes.

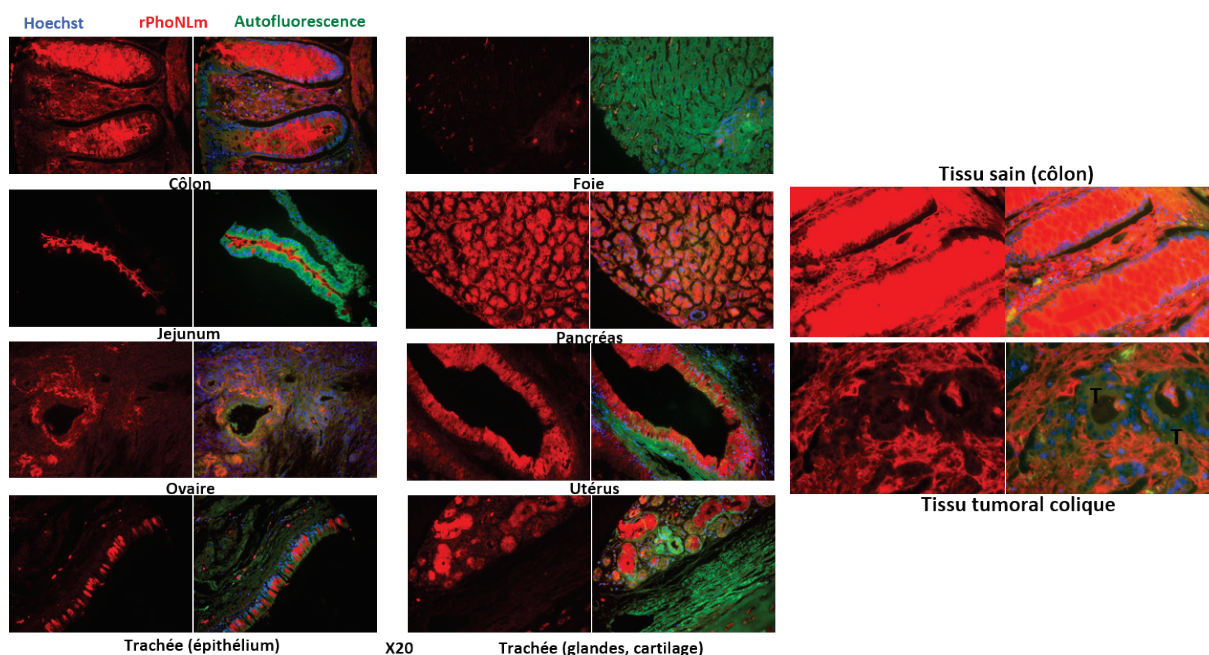


Figure IV-66. Marquage de tissus sains et comparaison de tissus sain et tumoraux coliques.

Utilisation de rPhoNLm à $20 \mu\text{g.mL}^{-1}$ marquées à la biotine avant l'ajout de streptavidine conjuguée à l'Alexa 549. Le marquage de rPhoNLm est en rouge, l'autofluorescence en vert et les noyaux en bleus. T= cellules tumorales.

4.3.2 Tests sur tumeurs induites chez le rat

Cette analyse sur tissus a été poursuivie sur des tumeurs induites chez le rat afin de comparer les marquages de rPhoSLm, rPhoNLm, rTrfbL2 et rTrfbL2-Y111A à l'IAB, Grenoble sous la supervision de Maxime Henry. Ces tumeurs ont été formées à partir de souches de cellules cancéreuses A549, issues de cellules épithéliales pulmonaires adénocarcinomiques, HT-29, dérivées de l'adénocarcinome du côlon, et U87, dérivé de glioblastome (cancer du cerveau). Des tissus de foie sains murins ont également été analysés.

Pour ces tests, rPhoSLm, rTrfbL2 et rTrfbL2-Y111A ont été marquées directement avec de la peroxydase (EZ-Link Plus Activated Peroxidase, ThermoFisher) contrairement à rPhoNLm qui a simplement été biotinylée. Dans ce dernier cas, rPhoNLm-biot a été incubée avec les tissus avant d'ajouter de la streptavidine-péroxydase permettant de révéler la fixation de la lectine. En effet, les premiers tests réalisés avec rPhoNLm marquée directement à la peroxydase n'ont montré aucune fixation de cette lectine sur les tissus, ce qui n'était pas cohérent avec les images obtenues précédemment sur les tissus humains. La peroxydase étant une protéine d'environ 44 kDa, il est probable qu'elle gêne l'interaction de rPhoNLm avec ses ligands lorsqu'elle n'est pas éloignée de la lectine par un bras biotine-streptavidine. La

péroxydase du kit EZ-Link Plus se fixe directement sur les amines primaires des protéines, c'est-à-dire sur les chaînes latérales des lysines et sur l'amine terminale. Le N-terminal est très proche du site d'interaction où les ligands se fixent dans les structures cristallographiques de rPhoNLm, et la présence d'une péroxydase à cet endroit pourrait gêner les interactions. Cependant la structure du N-terminal et la position de la lysine sont les mêmes chez rPhoSLm sans que cela ne gêne la reconnaissance. Les trois autres lysines de rPhoNLm sont positionnées sur chacune des grandes boucles de surface et sur un brin β de la dernière répétition. Dans l'hypothèse qu'un deuxième site d'interaction soit accessible en solution grâce au mouvement d'une des deux boucles, la présence d'une péroxydase pourrait là aussi gêner l'interaction. La réalité est sûrement bien plus complexe et doit prendre en compte aussi le fait que plusieurs péroxydases peuvent se fixer sur une même lectine. De ce fait, la distribution des amines primaires de rPhoNLm, pourrait conduire à une orientation des péroxydases défavorable aux interactions avec les ligands, contrairement à rPhoSLm.

Comme précédemment, les lectines ainsi marquées à $20\mu\text{g.mL}^{-1}$ ont été pré-incubées ou non avec leurs ligands respectifs (25 mM de chitobiose fucosylé en $\alpha 1,6$, 25 mM d'antigène TF, ou 100 mM GlcNAc) avant d'être mises en contact avec les tissus. Dans le cas de rPhoNLm-biot, les biotines endogènes des tissus ont été préalablement bloquées avec de la streptavidine (0.1 mg.mL^{-1}) puis de la biotine (0.5 mg.mL^{-1}).

4.3.2.1 rPhoSLm et rPhoNLm

Tout d'abord, une première comparaison a été faite pour le marquage de rPhoSLm et rPhoNLm sur les tumeurs A549, HT-29 et U87 et sur du foie sain (*Figure IV-67*). Les deux lectines reconnaissent toutes les coupes de tissus tumoraux testés, ce qui indiquerait la présence de core-fucose en surface. On observe également, comme cela avait été montré sur les tissus coliques humains, que le marquage de rPhoNLm est plus fort que celui de rPhoSLm. De plus, la distribution du marquage diffère entre les tissus : on observe un marquage extracellulaire des tumeurs A549 homogène sur toute la coupe, tandis que pour les coupes HT-29, certaines zones ne sont pas du tout marquées. Pour U87, le marquage forme des sillons caractéristiques qui pourraient correspondre aux cellules endothéliales. Le marquage est homogène sur toute la coupe comme pour A549. Concernant les coupes de foie sain, on n'observe aucun marquage avec rPhoSLm tandis qu'avec rPhoNLm, on peut voir un faible marquage de fond. Ces résultats semblent indiquer que la distribution de la core-fucosylation n'est pas toujours homogène selon les tissus. De plus, le foie sain murin semble avoir un faible taux de core-fucose.

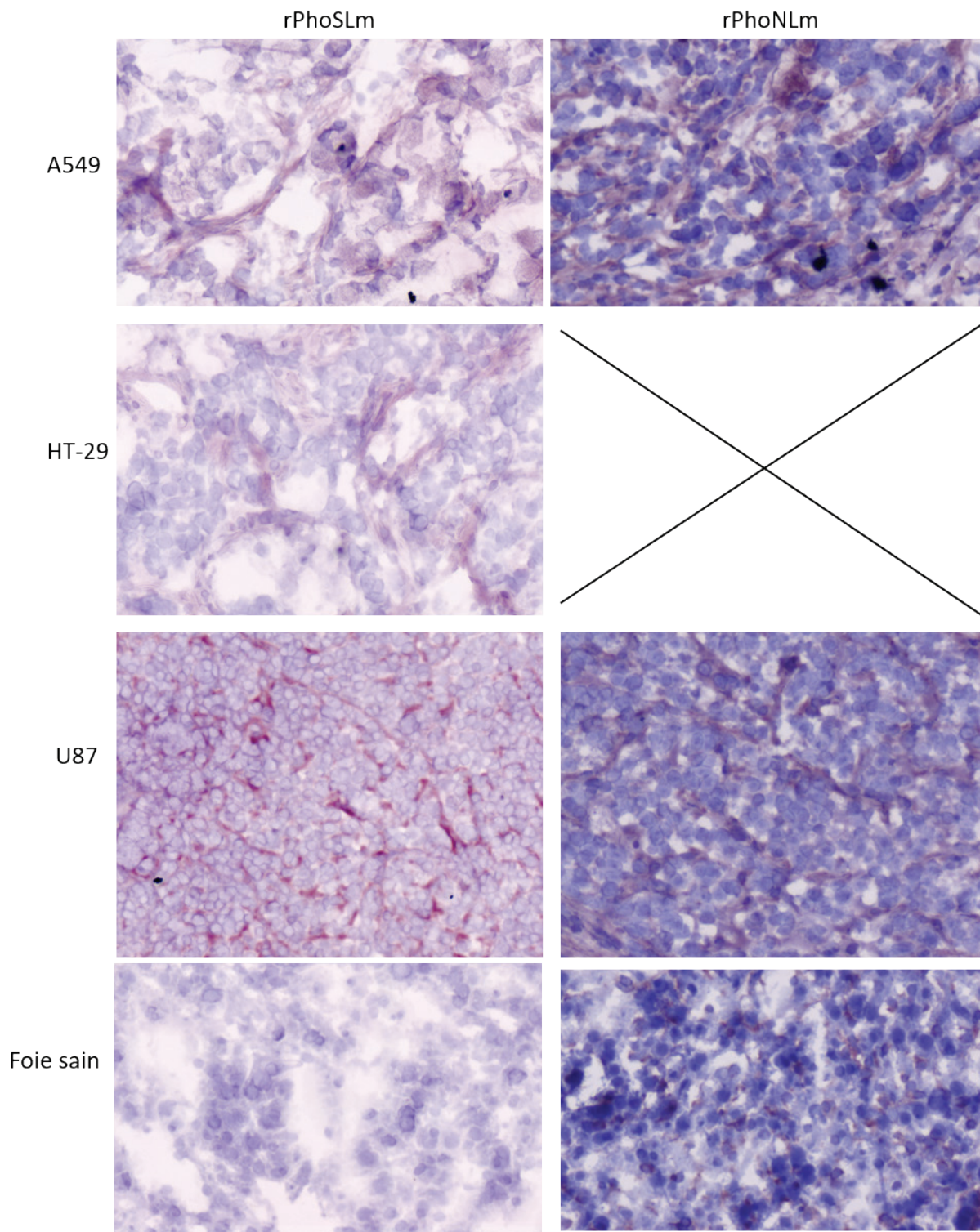


Figure IV-67. Comparaison du marquage de différents tissus par rPhoS_{Lm} et rPhoN_{Lm}.

Les lectines ont été utilisées à 20 µg.mL⁻¹. Marquage direct à la peroxydase (rPhoS_{Lm}) ou indirect avec de la biotine (rPhoN_{Lm}). Les lectines apparaissent en rouge/violet tandis que les noyaux sont colorés en bleu. Les noyaux des coupes de rPhoN_{Lm} ont été colorés plus longtemps et apparaissent donc en bleu plus intense. rPhoN_{Lm} n'a pas été testée sur les coupes HT-29. Les images ont été zoomées 10 X.

4.3.2.2 rTrfbL2 t rTrfbL2-Y111A

Ensuite, ces mêmes coupes de tissus ont été marquées avec rTrfbL2 et rTrfbL2-Y111A dans le but de comparer les capacités de marquage de ces deux lectines, et de comparer la distribution des marquages de l'antigène TF avec ceux de la core-fucosylation décrite dans le paragraphe précédent. Les résultats montrent que les deux lectines reconnaissent aussi toutes les coupes testées, suggérant la présence d'antigène TF en surface de ces tissus (**Figure IV-68**). On observe le même type de marquage entre rTrfbL2 et rTrfbL2-Y111A sur les différentes coupes avec des intensités similaires sauf pour les tumeurs A549 où l'intensité est plus faible avec le mutant Y111A. Par contre, comme précédemment avec rPhoS_{Lm} et rPhoN_{Lm}, le marquage est différent selon les tissus : on observe un marquage sous forme de tâches pour A549, un marquage de la matrice extracellulaire et des zones très intenses pour HT-29, un marquage des cellules endothéliales sous forme de sillons pour U87, et un faible marquage de fond pour le foie sain. On a donc un marquage similaire à rPhoS_{Lm} et rPhoN_{Lm} pour les coupes U87, mais une distribution différente des lectines pour les autres coupes tumorales, ce qui se traduirait par une distribution différente de l'antigène TF comparé à la core-fucosylation. Concernant le foie sain, là aussi l'antigène TF semble n'être que très peu présent comparé aux coupes de tumeurs. Une autre caractéristique est que pour toutes les coupes de tumeurs, le marquage n'est pas homogène mais beaucoup plus intense sur les extrémités par rapport au centre des coupes. Cela pourrait s'expliquer par le fait que les tumeurs grossissent par les extrémités et ont tendance à nécroser en leur centre, d'autant plus qu'on n'observe pas ce phénomène sur les coupes de foie sain. Cette particularité n'a pas été observée avec les lectines spécifiques de la core-fucosylation, et semble indiquer que l'antigène TF est surexprimé dans les zones de croissance tumorale et sous-exprimé dans les zones nécrosées.

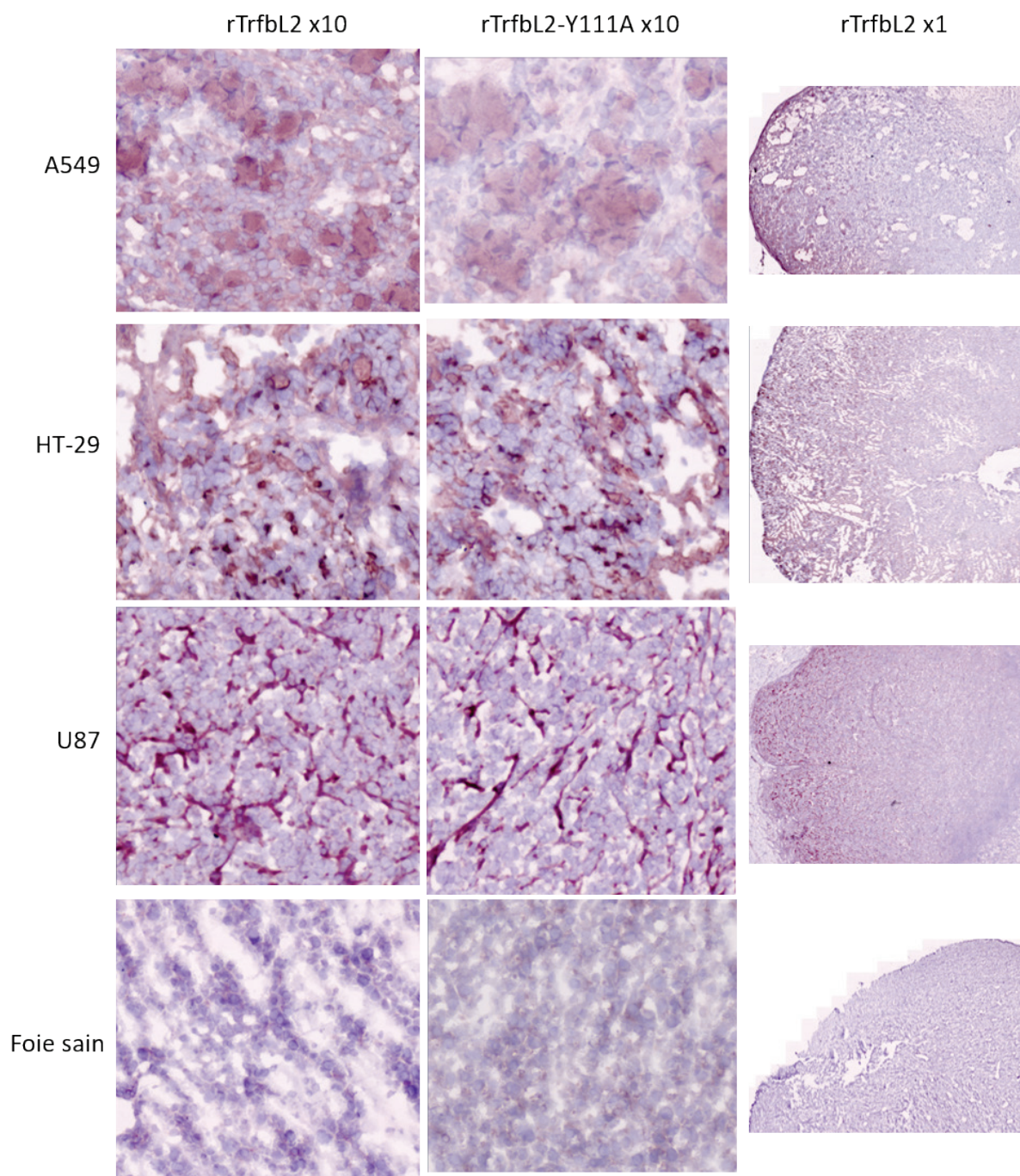


Figure IV-68. Comparaison du marquage de différents tissus par rTrfbL2 et rTrfbL2-Y111A.

Les lectines ont été utilisées à $20 \mu\text{g.mL}^{-1}$ et marquées directement à la peroxydase. Les lectines apparaissent en rouge/violet tandis que les noyaux sont colorés en bleu. Les images ont été zoomées 10 X quand cela est indiqué.

4.3.3 Tests d'inhibition

Dans le but d'analyser les spécificités de reconnaissance, des expériences d'inhibition avec les ligands spécifiques de chaque lectine ont été réalisées (*Figure IV-69*).

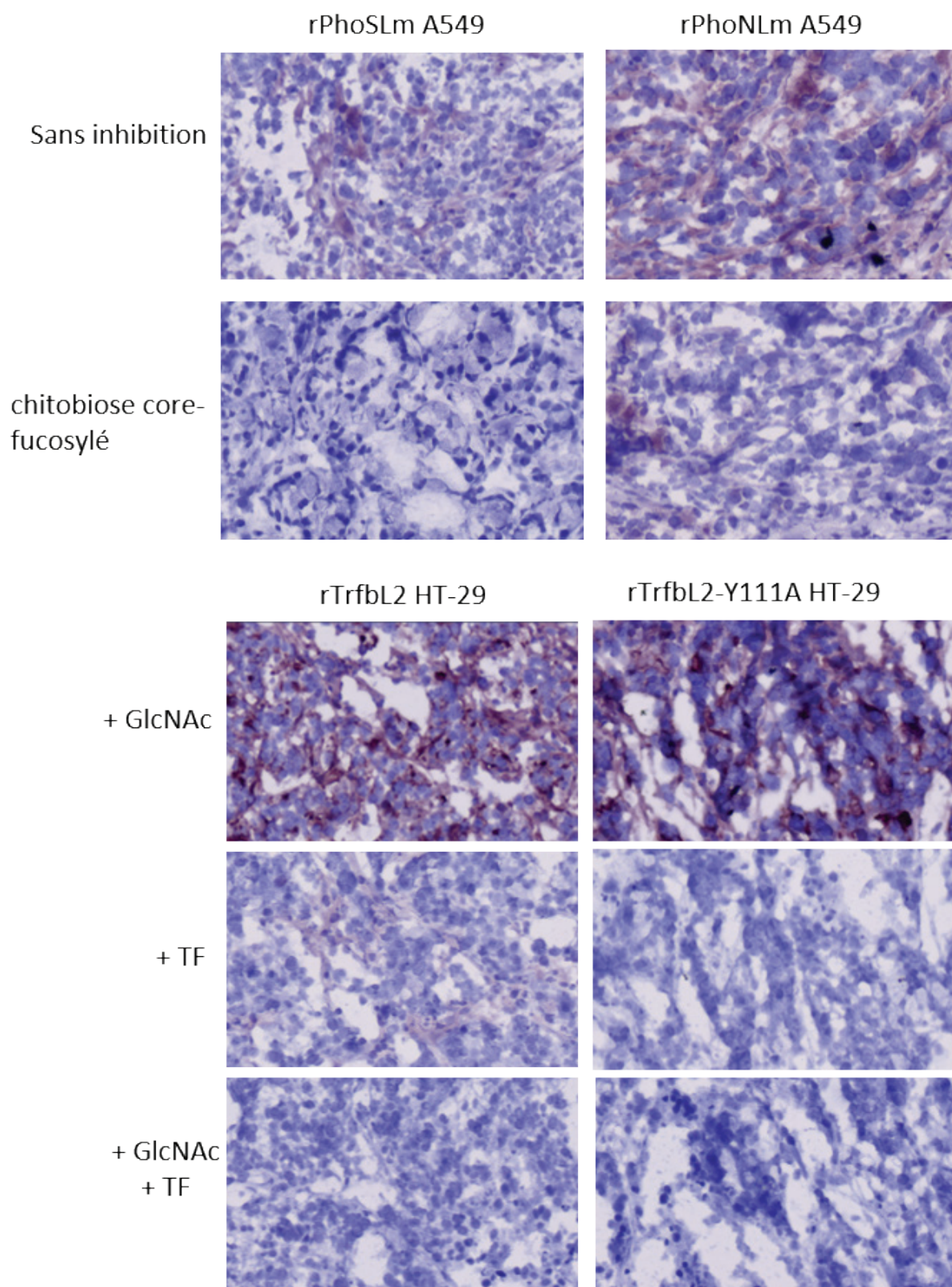


Figure IV-69. Tests d'inhibition de la reconnaissance des lectines rPhoS_LM, rPhoN_LM, rTrfbL2 et rTrfbL2-Y111A sur des coupes de tissus tumoraux.

Les lectines ont été utilisées à 20 µg.mL⁻¹, le chitobiose core-fucosylé et l'antigène TF à 25 mM et le GlcNAc à 100 mM. Marquage direct à la peroxydase (rPhoS_LM, rTrfbL2 et rTrfbL2-Y111A) ou indirecte avec de la biotine (rPhoN_LM).

Ces tests d'inhibition tendent à confirmer la reconnaissance spécifique de la core-fucosylation par rPhoS_LM et rPhoN_LM sur les coupes A549 avec une extinction complète du signal pour rPhoS_LM et une forte diminution pour rPhoN_LM. L'extinction du signal n'est pas complète pour rPhoN_LM car cette lectine marque plus fortement les tissus. Il aurait fallu diminuer la concentration en lectine ou augmenter la concentration d'inhibiteur. Ces tests tendent également à confirmer la reconnaissance spécifique de l'antigène TF par rTrfbL2 et rTrfbL2-Y111A sur les coupes HT-29 avec une forte diminution ou une extinction complète du signal pour rTrfbL2 et rTrfbL2-Y111A respectivement. Il est intéressant de remarquer que l'ajout de GlcNAc en plus de l'antigène TF est nécessaire pour supprimer totalement les interactions de rTrfbL2 sur les coupes HT-29 alors que ce n'est pas le cas du mutant Y111A. Cela semble montrer que la mutation Y111A suffit à supprimer les interactions avec le GlcNAc sur ces coupes. Pour les deux lectines, l'ajout de GlcNAc seul n'a pas d'impact significatif sur le marquage.

4.4 Conclusions générales de l'étude biologique sur cellules et tissus

Toutes les lectines testées semblent être capables de reconnaître spécifiquement leurs ligands sur les cellules et tissus utilisés dans cette étude. Elles n'ont cependant pas toutes les mêmes capacités. Ainsi, rPhoS_LM a montré le plus fort potentiel pour la reconnaissance de la core-fucosylation sur cellules, tandis que sur tissus, c'est rPhoN_LM qui apparaît la plus efficace. Concernant la reconnaissance de l'antigène TF, sur cellules, rTrfbL2 et rTrfbL2-Y111A semblent fixer toutes les deux spécifiquement l'antigène TF, sans qu'il n'y ait de biais apparent avec la reconnaissance du GlcNAc. Cependant, sur tissu, on observe une faible reconnaissance du GlcNAc par rTrfbL2, contrairement au mutant rTrfbL2-Y111A qui semble reconnaître uniquement l'antigène TF et est donc plus adapté aux études de lectine-histochimie.

Ces résultats prouvent encore une fois l'intérêt de ces lectines pour la reconnaissance de marqueurs spécifiques. Actuellement, d'autres lectines sont utilisées pour la reconnaissance de la core-fucosylation comme AAL, AOL, PSA et LCA par exemple. Une étude comparative a été réalisée en 2009 avec ces lectines et a montré que AAL et AOL ne sont pas spécifiques du core fucose puisqu'elles reconnaissent aussi le fucose lié en α 1-2, α 1-3, et α 1-4, que l'on retrouve sur les structures Lewis et les groupes sanguins (Tateno et al., 2009). L'article met en avant l'intérêt de PSA et LCA qui ont davantage de spécificité pour le core-fucose comparé aux autres types de fucosylation, bien que la fixation de ces deux lectines analysée en FACS soit partiellement inhibée par du Met- α Man car elles reconnaissent les oligosaccharides α -

mannosylés. Malgré tout plusieurs études ont utilisé LCA pour la reconnaissance spécifique de la core fucosylation en FACS (Tateno et al., 2009; Tu et al., 2017; ou Wang et al., 2006) et en lectine-histochimie (Li et al., 2016). rPhoS_Lm et rPhoN_Lm, de par leur stricte reconnaissance de la core-fucosylation pourraient être des outils encore plus spécifiques et leur utilisation en parallèle d'autres lectines tel que LCA ou AAL, permettrait d'affiner les analyses lorsque plusieurs types de fucosylations sont présents.

Concernant la reconnaissance de l'antigène TF, les autres lectines de la même famille disponibles (BEL, ABL, SRL, XCL) ont une double spécificité pour le GlcNAc et ne sont pas toujours strictement spécifiques de l'antigène TF. Les résultats obtenus avec rTrfbL2 et rTrfbL2-Y111A sont encourageants pour l'utilisation de ces lectines dans la reconnaissance de l'antigène TF, sans biais de fixation avec le GlcNAc.

V.

**CONCLUSIONS ET
PERSPECTIVES**

Ce travail a porté sur le développement de lectines pour la reconnaissance spécifique de la core-fucosylation et de l'antigène TF, deux marqueurs que l'on retrouve à des taux anormaux dans de nombreux cancers. Trois lectines ont été étudiées dans cet objectif: rPhoSL et rPhoNL pour la reconnaissance de la core-fucosylation et rTrfbL2 pour la reconnaissance de l'antigène TF.

La lectine rTrfbL2 produite sous forme recombinante dans *E. coli* a des propriétés prometteuses pour la reconnaissance de l'antigène TF. C'est une protéine très stable, qui a deux sites d'interaction distincts pour le GalNAc et pour le GlcNAc et dont l'affinité pour l'antigène TF est de l'ordre du micromolaire. Elle ne reconnaît pas le GalNAc seul ni l'antigène Tn, ce qui est un gage de sa spécificité. Elle a cependant une faible affinité de l'ordre du millimolaire pour le GlcNAc et des mutants de cette lectine ont donc été élaborés dans le but d'invalider le site de fixation du GlcNAc afin d'obtenir une lectine capable de reconnaître uniquement l'antigène TF. Parmi les cinq mutants réalisés, seuls deux d'entre eux (rTrfbL2-Y111A et rTrfbL2-Y111Q) ont pu être correctement purifiés, et le mutant Y111A a pu faire l'objet d'une analyse biochimique et structurale complète. Ce mutant a la même affinité pour l'antigène TF que la protéine sauvage et a une affinité diminuée pour le GlcNAc qui n'est plus mesurable en microcalorimétrie, mais permet toujours de purifier la lectine sur une colonne d'affinité au GlcNAc. Cela est probablement dû à la multivalence de rTrfbL2 qui, en formant des tétramères, présente quatre sites d'interaction par molécule ce qui crée un phénomène d'avidité sur la résine qui compense la faible affinité. La résolution de la structure de ce mutant a permis de visualiser l'impact de la mutation et notamment la perte de l'interaction avec le résidu 111.

Pour les autres mutants visant à diminuer encore davantage l'affinité de cette lectine pour le GlcNAc, des nombreuses difficultés de purification ont été rencontrées et sont probablement liées à l'instabilité des protéines mutantes. L'analyse de la structure a permis d'ailleurs de mieux comprendre en quoi les mutations réalisées pouvaient déstabiliser la structure à cause de la perte d'interactions entre certains acides aminés. Ces problèmes d'instabilité n'ayant pas pu être surmontés, il serait intéressant à l'avenir de pouvoir réaliser des études de modélisation moléculaire afin de définir d'autres mutations moins destabilisantes permettant d'abroger la fixation du GlcNAc. Cette étude n'ayant pas pu être réalisée dans le cadre de ce travail, seul le mutant Y111A a ensuite été utilisé pour les études sur cellules et tissus cancéreux.

rTrfbL2 et rTrfbL2-Y111A ont montré des résultats encourageants pour le marquage spécifique de l'antigène TF sur les cellules cancéreuses testées, et étonnamment, sans qu'il n'y ait de biais apparent avec la reconnaissance du GlcNAc pour les deux lectines. Sur tissus par contre, on observe un marquage dû au GlcNAc pour rTrfbL2 qui n'est pas visible avec le mutant Y111A. rTrfbL2-Y111A semble donc avoir une affinité suffisamment faible pour le GlcNAc pour être utilisée pour le marquage spécifique de l'antigène TF sur cellules et tissus. C'est un outil qui pourra trouver des applications en diagnostic de cancers par exemple, et qui contrairement aux autres lectines de la famille de rTrfbL2, ne reconnaît quasiment plus le GlcNAc. Cette mutation semble être suffisante pour faire de rTrfbL2-Y111A un bon outil.

Les lectines rPhoSL et rPhoNL sont des lectines homologues strictement spécifiques de la core-fucosylation et particulièrement stables bien que rPhoSL ait une meilleure affinité pour ses ligands que rPhoNL. L'analyse structurale de rPhoSL et rPhoNL a permis de découvrir que ces lectines avaient un nouveau repliement que l'on a appelé β -prisme III et l'analyse des interactions de rPhoSL avec des ligands core-fucosylés à l'échelle atomique a permis de définir une séquence consensus caractéristique de cette nouvelle famille de lectines. Les acides aminés impliqués dans l'interaction étant alors connus, la comparaison entre les séquences de rPhoSL et de rPhoNL a ensuite révélé que les sites d'interactions de rPhoNL n'étaient pas optimaux, ce qui a permis de comprendre d'où venait la très faible affinité de rPhoNL pour ses ligands. Les structures des deux lectines ont également permis de suspecter les premiers acides aminés N-terminaux, ajoutés lors du clonage, de diminuer l'affinité de ces protéines, qui même pour rPhoSL est 100 fois plus faible que pour la lectine native.

Les mutants rPhoSLm et rPhoNLm ont été clonés, produits et purifiés avec succès. Après la suppression des deux acides aminés N-terminaux et l'optimisation de la séquence de rPhoNL, ces mutants ont la même affinité pour les ligands core-fucosylés qui est 10 fois supérieure à celle mesurée précédemment pour rPhoSL. rPhoNLm conserve cependant d'autres limites qui sont dues à sa structure. En effet, tandis que rPhoSL possède 3 sites d'interaction fonctionnels pour la lectine trimérique, un seul interagit avec certitude avec les sucres chez rPhoNL à cause de deux grandes boucles de surface qui recouvrent les deux autres sites d'interaction. Bien que rPhoNLm forme des dimères dans le cristal, elle est monomérique en solution, ce qui indiquerait qu'au moins une des deux boucles est capable de bouger pour permettre l'insertion d'un ligand puisque sa capacité d'hémagglutination nécessite qu'elle soit multivalente. Une étude a été réalisée pour tenter de créer un dimère stable de rPhoNLm mais cela n'a pas abouti.

Le potentiel des lectines rPhoSLm et rPhoNLM pour la reconnaissance spécifique de la core-fucosylation sur cellules et tissus biologiques a été analysé. Ce sont des outils performants dont la stricte spécificité est un plus comparé aux autres lectines reconnaissant le core-fucose et dont certaines sont actuellement utilisées dans la recherche et le diagnostic appliqués aux cancers.

Ce travail pourra par la suite être poursuivi pour y apporter de nouveaux éléments et optimisations. Tout d'abord, les rendements de production et purification de rPhoSL et rPhoNL (1 et 2 mg par litre de culture respectivement) mériteraient d'être améliorés. Il a été montré que pour améliorer le rendement des protéines exprimées dans le périplasma, il était bénéfique d'ajuster les niveaux d'expression du gène cible afin de dé-saturer le translocon permettant à la protéine de passer du cytoplasme au périplasma (Schlegel et al., 2013). D'autres facteurs peuvent être pris en compte comme le milieu de culture, le choix du plasmide et du promoteur, l'utilisation de la voie SRP ou Sec pour la translocation, la surexpression des protéines Dsb qui catalysent la formation des ponts disulfures, ou encore la co-expression de chaperonnes pour aider au repliement (de Marco, 2009).

La séquence des mutants rPhoSLm et rPhoNLM pourrait aussi être optimisée pour augmenter encore davantage l'affinité de ces lectines. En particulier, la mutation de la phénylalanine 23 en tyrosine permettrait de nouvelles interactions hydrogènes avec le GlcNAc1 du noyau des N-glycanes. De plus, chez rPhoNLM, la proline 35 pourra être mutée en asparagine pour favoriser les interactions avec la branche α 1,6 des N-glycanes comme observé chez rPhoSL. Toujours avec rPhoNLM, il pourrait être intéressant de poursuivre les essais pour créer un dimère stable en augmentant les interactions à l'interface de dimérisation. De plus, pour optimiser la formation des ponts disulfures reliant deux monomères, rPhoNLMSS pourrait être produite dans un système optimisé comme évoqué précédemment (surexpression de DsbA et DsbC, autre voie de translocation...). Une autre possibilité pour augmenter le nombre de sites d'interaction actifs de rPhoNLM est de modifier la séquence des grandes boucles qui recouvrent deux des trois sites d'interaction pour les rendre plus flexibles. Cela pourra être réalisé en s'appuyant sur les séquences des boucles de la lectine homologue SL2-1, dont deux sites d'interaction sont accessibles. Ces modifications de rPhoNLM pourraient apporter des propriétés intéressantes de reconnaissance des ligands sur une surface avec une organisation et/ou un nombre de sites d'interaction différents de rPhoSLm.

Les deux lectines rPhoSLm et rPhoNLM pourront également être testées sur des anticorps thérapeutiques, core-fucosylés ou non dans le but d'évaluer leur capacité pour le contrôle

qualité de ces anticorps. Si ces tests s'avèrent prometteurs, ces lectines pourront être greffées sur une résine pour créer une colonne à lectines permettant de simplifier la séparation des anticorps non core-fucosylés. Elles pourraient aussi permettre l'analyse des clones d'anticorps lors de leur production et la sélection de ceux qui produisent des anticorps non fucosylés. L'absence de core-fucosylation étant associée à une meilleure efficacité de certains anticorps thérapeutiques, ces lectines pourront ainsi contribuer à la diminution des doses des traitements administrés et donc de leurs coûts.

Concernant les trois lectines rTrfbL2-Y111A, rPhoSLm, et rPhoNLm, leur capacité de marquage sur cellules et tissus pourra être analysée avec davantage de précision, sur une plus grande variété de cellules et tissus. Notamment, la comparaison des marquages entre tissus sains et cancéreux humains est un point important qui n'a pas pu être suffisamment exploité au cours de ce travail par manque de matériel et les difficultés d'accès aux coupes de tissus. De même, la comparaison des marquages avec d'autres lectines reconnaissant le même type de marqueur de cancer permettrait d'évaluer plus précisément l'atout de ces nouveaux outils pour le diagnostic de cancers. Pour rPhoSLm et rPhoNLm, il serait aussi intéressant de tester ces lectines dans l'analyse de sérums de patients puisque certains cancers sont associés à une augmentation de la core-fucosylation de glycoprotéines spécifiques (AFP-L3, PSA, Hpt). Cela peut être réalisé sur le principe d'un test ELISA-sandwich comme cela a déjà été réalisé avec PhoSL native par exemple (Kobayashi et al., 2012).

Enfin, ces lectines pourraient être intégrées dans une puce à lectines pour le diagnostic et le suivi de cancers ou pour le contrôle qualité ou l'analyse de la glycosylation sur différents types d'échantillons que ce soit dans le domaine fondamental (glycomique) ou biomédical. Cela a par exemple été réalisé pour le cancer du sein avec une puce contenant 45 lectines de spécificités variables pour des N- et des O-glycanes (Fry et al., 2011). Cette méthode permet d'analyser des tissus, du sérum ou encore des urines en faible quantité. Il s'agit d'un outil très prometteur qui permet une analyse simple, rapide et détaillée de la glycosylation (Pilobello et al., 2005). L'intérêt est ici d'avoir des protéines recombinantes, hautement reproductibles et de qualité définie, ce qui n'est pas le cas actuellement des puces disponibles où seules des lectines extraites et commerciales sont retrouvées.

La découverte de nouvelles lectines spécifiques des différents marqueurs de cancers connus (core-fucosylation, troncation des O-glycanes, sialylation...) contribue à enrichir les outils d'analyse du glycome, mieux comprendre et diagnostiquer des maladies telles que les cancers et aider la recherche en glycobiologie et biotechnologie.

VI.

BIBLIOGRAPHIE

A

Agirre, J., Iglesias-Fernández, J., Rovira, C., Davies, G.J., Wilson, K.S., and Cowtan, K.D. (2015). Privateer: software for the conformational validation of carbohydrate structures. *Nat. Struct. Mol. Biol.* *22*, 833.

Agrawal, P., Fontanals-Cirera, B., Sokolova, E., Jacob, S., Vaiana, C.A., Argibay, D., Davalos, V., McDermott, M., Nayak, S., Darvishian, F., et al. (2017). A Systems Biology Approach Identifies FUT8 as a Driver of Melanoma Metastasis. *Cancer Cell* *31*, 804–819.e7.

Aich, U., Liu, X., Rao, S., and Pohl, C. (2013). Separation of 2AB Labeled N-Glycans from Bovine Fetuin on a Novel Mixed-Mode Stationary Phase. *Thermo Sci. Appl. Note* 20754.

Al Teneiji, A., Bruun, T.U.J., Sidky, S., Cordeiro, D., Cohn, R.D., Mendoza-Londono, R., Moharir, M., Raiman, J., Siriwardena, K., Kyriakopoulou, L., et al. (2017). Phenotypic and genotypic spectrum of congenital disorders of glycosylation type I and type II. *Mol. Genet. Metab.* *120*, 235–242.

Aoki-Kinoshita, K., Agravat, S., Aoki, N.P., Arpinar, S., Cummings, R.D., Fujita, A., Fujita, N., Hart, G.M., Haslam, S.M., Kawasaki, T., et al. (2016). GlyTouCan 1.0 – The international glycan structure repository. *Nucleic Acids Res.* *44*, D1237–D1242.

Audfray, A., Beldjoudi, M., Breiman, A., Hurbin, A., Boos, I., Unverzagt, C., Bouras, M., Lantuejoul, S., Coll, J.-L., Varrot, A., et al. (2015). A Recombinant Fungal Lectin for Labeling Truncated Glycans on Human Cancer Cells. *PLOS ONE* *10*, e0128190.

B

Barkai-Golan, R., Mirelman, D., and Sharon, N. (1978). Studies on growth inhibition by lectins of *penicillia* and *aspergilli*. *Arch. Microbiol.* *116*, 119–124.

Becker, D.J., and Lowe, J.B. (2003). Fucose: biosynthesis and biological function in mammals. *Glycobiology* *13*, 41R–53R.

Berkel, P.H.C. van, Gerritsen, J., Perdok, G., Valbjørn, J., Vink, T., Winkel, J.G.J. van de, and Parren, P.W.H.I. (2009). N-linked glycosylation is an important parameter for optimal selection of cell lines producing biopharmaceutical human IgG. *Biotechnol. Prog.* *25*, 244–251.

Berkmen, M. (2012). Production of disulfide-bonded proteins in *Escherichia coli*. *Protein Expr. Purif.* *82*, 240–251.

Birck, C., Damian, L., Marty-Detraves, C., Lougarre, A., Schulze-Briese, C., Koehl, P., Fournier, D., Paquereau, L., and Samama, J.-P. (2004). A New Lectin Family with Structure Similarity to Actinoporins Revealed by the Crystal Structure of *Xerocomus chrysenteron* Lectin XCL. *J. Mol. Biol.* *344*, 1409–1420.

den Blaauwen, T., and Driessen, A.J.M. (1996). Sec-dependent preprotein translocation in bacteria. *Arch. Microbiol.* *165*, 1–8.

Bojarová, P., and Křen, V. (2016). Sugared biomaterial binding lectins: achievements and perspectives. *Biomater. Sci.* *4*, 1142–1160.

Bonnardel, F., Mariethoz, J., Salentin, S., Robin, X., Schroeder, M., Perez, S., Lisacek, F., and Imberty, A. (2019). UniLectin3D, a database of carbohydrate binding proteins with curated information on 3D structures and interacting ligands. *Nucleic Acids Res.* *47*, D1236–D1244.

Bovi, M., Carrizo, M.E., Capaldi, S., Perduca, M., Chiarelli, L.R., Galliano, M., and Monaco, H.L. (2011). Structure of a lectin with antitumoral properties in king bolete (*Boletus edulis*) mushrooms. *Glycobiology* 21, 1000–1009.

Bresalier, R.S., Byrd, J.C., Wang, L., and Raz, A. (1996). Colon Cancer Mucin: A New Ligand for the β -Galactoside-binding Protein Galectin-3. *Cancer Res.* 56, 4354–4357.

Butcher, L.D., Garcia, M., Arnold, M., Ueno, H., Goel, A., and Boland, C.R. (2014). Immune response to JC virus T antigen in patients with and without colorectal neoplasia. *Gut Microbes* 5, 468–475.

C

Calvete, J.J., Costa, F.H., Saker-Sampaio, S., Murciano, M.P., Nagano, C.S., Cavada, B.S., Grangeiro, T.B., Ramos, M.V., Bloch, C., Silveira, S.B., et al. (2000). The amino acid sequence of the agglutinin isolated from the red marine alga *Bryothamnion triquetrum* defines a novel lectin structure. *Cell. Mol. Life Sci. CMLS* 57, 343–350.

Carrizo, M.E., Capaldi, S., Perduca, M., Irazoqui, F.J., Nores, G.A., and Monaco, H.L. (2005). The Antineoplastic Lectin of the Common Edible Mushroom (*Agaricus bisporus*) Has Two Binding Sites, Each Specific for a Different Configuration at a Single Epimeric Hydroxyl. *J. Biol. Chem.* 280, 10614–10623.

Ceroni, A., Passerini, A., Vullo, A., and Frasconi, P. (2006). DISULFIND: a disulfide bonding state and cysteine connectivity prediction server. *Nucleic Acids Res.* 34, W177–181.

Ceroni, A., Maass, K., Geyer, H., Geyer, R., Dell, A., and Haslam, S.M. (2008). GlycoWorkbench: a tool for the computer-assisted annotation of mass spectra of glycans. *J. Proteome Res.* 7, 1650–1659.

Chachadi, V.B., Inamdar, S.R., Yu, L.-G., Rhodes, J.M., and Swamy, B.M. (2011). Exquisite binding specificity of *Sclerotium rolfii* lectin toward TF-related O-linked mucin-type glycans. *Glycoconj. J.* 28, 49–56.

Chang, A., Singh, S., Phillips, G.N., and Thorson, J.S. (2011). Glycosyltransferase structural biology and its role in the design of catalysts for glycosylation. *Curr. Opin. Biotechnol.* 22, 800–808.

Chen, K., Gentry-Maharaj, A., Burnell, M., Steentoft, C., Marcos-Silva, L., Mandel, U., Jacobs, I., Dawnay, A., Menon, U., and Blixt, O. (2013). Microarray Glycoprofiling of CA125 Improves Differential Diagnosis of Ovarian Cancer. *J. Proteome Res.* 12, 1408–1418.

Choi, S.H., Lyu, S.Y., and Park, W.B. (2004). Mistletoe lectin induces apoptosis and telomerase inhibition in human A253 cancer cells through dephosphorylation of akt. *Arch. Pharm. Res.* 27, 68.

Chojnacki, S., Cowley, A., Lee, J., Foix, A., and Lopez, R. (2017). Programmatic access to bioinformatics tools from EMBL-EBI update: 2017. *Nucleic Acids Res.* 45, W550–W553.

Chung, C.-Y., Majewska, N.I., Wang, Q., Paul, J.T., and Betenbaugh, M.J. (2017). SnapShot: N-glycosylation processing pathways across kingdoms. *Cell* 171, 258–258.

Cregg, J.M., Cereghino, J.L., Shi, J., and Higgins, D.R. (2000). Recombinant protein expression in *Pichia pastoris*. *Mol. Biotechnol.* 16, 23–52.

Croci, D.O., Cerliani, J.P., Dalotto-Moreno, T., Méndez-Huergo, S.P., Mascanfroni, I.D., Dergan-Dylon, S., Toscano, M.A., Caramelo, J.J., García-Vallejo, J.J., Ouyang, J., et al. (2014). Glycosylation-Dependent Lectin-Receptor Interactions Preserve Angiogenesis in Anti-VEGF Refractory Tumors. *Cell* 156, 744–758.

D

Dam, T.K., Talaga, M.L., Fan, N., and Brewer, C.F. (2016). Chapter Four - Measuring Multivalent Binding Interactions by Isothermal Titration Calorimetry. In *Methods in Enzymology*, A.L. Feig, ed. (Academic Press), pp. 71–95.

Damian, L., Fournier, D., Winterhalter, M., and Paquereau, L. (2005). Determination of thermodynamic parameters of *Xerocomus chrysenteron* lectin interactions with N-acetylgalactosamine and Thomsen-Friedenreich antigen by isothermal titration calorimetry. *BMC Biochem.* 6, 11.

Dan, X., Liu, W., and Ng, T.B. (2016). Development and Applications of Lectins as Biological Tools in Biomedical Research. *Med. Res. Rev.* 36, 221–247.

daSilva, L.L.P., Molfetta-Machado, J.B. de, Panunto-Castelo, A., Denecke, J., Goldman, G.H., Roque-Barreira, M.-C., and Goldman, M.H.S. (2005). cDNA cloning and functional expression of KM⁺, the mannose-binding lectin from *Artocarpus integrifolia* seeds. *Biochim. Biophys. Acta BBA - Gen. Subj.* 1726, 251–260.

Debarbieux, L., and Beckwith, J. (1998). The reductive enzyme thioredoxin 1 acts as an oxidant when it is exported to the *Escherichia coli* periplasm. *Proc. Natl. Acad. Sci.* 95, 10751–10756.

Demain, A.L., and Vaishnav, P. (2009). Production of recombinant proteins by microbes and higher organisms. *Biotechnol. Adv.* 27, 297–306.

Doig, A.J., and Williams, D.H. (1991). Is the hydrophobic effect stabilizing or destabilizing in proteins?: The contribution of disulphide bonds to protein stability. *J. Mol. Biol.* 217, 389–398.

Dombkowski, A.A. (2003). Disulfide by DesignTM: a computational method for the rational design of disulfide bonds in proteins. *Bioinformatics* 19, 1852–1853.

Donnelly, E.H., and Goldstein, I.J. (1970). Glutaraldehyde-insolubilized concanavalin A: an adsorbent for the specific isolation of polysaccharides and glycoproteins. *Biochem. J.* 118, 679–680.

Doublié, S. (1997). [29] Preparation of selenomethionyl proteins for phase determination. In *Methods in Enzymology*, (Academic Press), pp. 523–530.

E

Ecker, D.M., Jones, S.D., and Levine, H.L. (2014). The therapeutic monoclonal antibody market. *MAbs* 7, 9–14.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of *Coot*. *Acta Crystallogr. D Biol. Crystallogr.* 66, 486–501.

F

Falconer, R., Errington, R., Shnyder, S., Smith, P., and Patterson, L. (2012). Polysialyltransferase: A New Target in Metastatic Cancer. *Curr. Cancer Drug Targets* 12, 925–939.

Fox, N.K., Brenner, S.E., and Chandonia, J.-M. (2014). SCOPe: Structural Classification of Proteins—extended, integrating SCOP and ASTRAL data and classification of new structures. *Nucleic Acids Res.* 42, D304–D309.

Fry, S.A., Afrough, B., Lomax-Browne, H.J., Timms, J.F., Velentzis, L.S., and Leathem, A.J. (2011). Lectin microarray profiling of metastatic breast cancers. *Glycobiology* 21, 1060–1070.

Fujii, H., Shinzaki, S., Iijima, H., Wakamatsu, K., Iwamoto, C., Sobajima, T., Kuwahara, R., Hiyama, S., Hayashi, Y., Takamatsu, S., et al. (2016). Core Fucosylation on T Cells, Required for Activation of T-Cell Receptor Signaling and Induction of Colitis in Mice, Is Increased in Patients With Inflammatory Bowel Disease. *Gastroenterology* 150, 1620–1632.

Fukumori, F., Takeuchi, N., Hagiwara, T., Ohbayashi, H., Endo, T., Kochibe, N., Nagata, Y., and Kobata, A. (1990). Primary Structure of a Fucose-Specific Lectin Obtained from a Mushroom, *Aleuria aurantia*. *J. Biochem. (Tokyo)* 107, 190–196.

G

Gabius, H.-J. (2018). The sugar code: Why glycans are so important. *Biosystems* 164, 102–111.

Gallagher, S.R. (2007). One-Dimensional SDS Gel Electrophoresis of Proteins. *Curr. Protoc. Cell Biol.* 37, 6.1.1-6.1.38.

Gemeiner, P., Mislovičová, D., Tkáč, J., Švitel, J., Pätoprstý, V., Hrabárová, E., Kogan, G., and Kožár, T. (2009). Lectinomics: II. A highway to biomedical/clinical diagnostics. *Biotechnol. Adv.* 27, 1–15.

Glinsky, V.V., Huflejt, M.E., Glinsky, G.V., Deutscher, S.L., and Quinn, T.P. (2000). Effects of Thomsen-Friedenreich Antigen-specific Peptide P-30 on β -Galactoside-mediated Homotypic Aggregation and Adhesion to the Endothelium of MDA-MB-435 Human Breast Carcinoma Cells. *Cancer Res.* 60, 2584–2588.

Gomes Ferreira, I., Pucci, M., Venturi, G., Malagolini, N., Chiricolo, M., and Dall’Olio, F. (2018). Glycosylation as a Main Regulator of Growth and Death Factor Receptors Signaling. *Int. J. Mol. Sci.* 19, 580.

H

Hakomori, S. (2002). Glycosylation defining cancer malignancy: new wine in an old bottle. *Proc. Natl. Acad. Sci.* 99, 10231–10233.

Hart, G.W., and Copeland, R.J. (2010). Glycomics Hits the Big Time. *Cell* 143, 672–676.

Hashim, O.H., Jayapalan, J.J., and Lee, C.-S. (2017). Lectins: an effective tool for screening of potential cancer biomarkers. *PeerJ* 5, e3784.

Hassan, M.A.A., Rouf, R., Tiralongo, E., May, T.W., and Tiralongo, J. (2015). Mushroom Lectins: Specificity, Structure and Bioactivity Relevant to Human Disease. *Int. J. Mol. Sci.* 16, 7802–7838.

Hayes, J.H., and Barry, M.J. (2014). Screening for Prostate Cancer With the Prostate-Specific Antigen Test: A Review of Current Evidence. *JAMA* 311, 1143–1149.

Helenius, A., and Aebi, M. (2001). Intracellular functions of N-linked glycans. *Science* 291, 2364–2369.

Hennet, T., and Cabalzar, J. (2015). Congenital disorders of glycosylation: a concise chart of glycoalkaloid dysfunction. *Trends Biochem. Sci.* 40, 377–384.

Henry, N.L., and Hayes, D.F. (2012). Cancer biomarkers. *Mol. Oncol.* 6, 140–146.

Hester, G., Kaku, H., Goldstein, I.J., and Wright, C.S. (1995). Structure of mannan-specific snowdrop (*Galanthus nivalis*) lectin is representative of a new plant lectin family. *Nat. Struct. Mol. Biol.* 2, 472–479.

Holm, L., and Rosenström, P. (2010). Dali server: conservation mapping in 3D. *Nucleic Acids Res.* 38, W545–W549.

Holmgren, A. (1972). Tryptophan fluorescence study of conformational transitions of the oxidized and reduced form of thioredoxin. *J. Biol. Chem.* 247, 1992–1998.

Hung, J.-S., Huang, J., Lin, Y.-C., Huang, M.-J., Lee, P.-H., Lai, H.-S., Liang, J.-T., and Huang, M.-C. (2014). C1GALT1 overexpression promotes the invasive behavior of colon cancer cells through modifying O-glycosylation of FGFR2. *Oncotarget* 5, 2096–2106.

I

Inamdar, S.R., Savanur, M.A., Eligar, S.M., Chachadi, V.B., Nagre, N.N., Chen, C., Barclays, M., Ingle, A., Mahajan, P., Borges, A., et al. (2012). The TF-antigen binding lectin from *Sclerotium rolfsii* inhibits growth of human colon cancer cells by inducing apoptosis in vitro and suppresses tumor growth in vivo. *Glycobiology* 22, 1227–1235.

Inamdar, S.R., Eligar, S.M., Ballal, S., Belur, S., Kalraiya, R.D., and Swamy, B.M. (2016). Exquisite specificity of mitogenic lectin from *Cephalosporium curvulum* to core fucosylated N-glycans. *Glycoconj. J.* 33, 19–28.

J

Jaeken, J., and Péanne, R. (2017). What is new in CDG? *J. Inher. Metab. Dis.* 40, 569–586.

Janzen, D.H., Juster, H.B., and Liener, I.E. (1976). Insecticidal action of the phytohemagglutinin in black beans on a bruchid beetle. *Science* 192, 795–796.

Jefferis, R., Lund, J., and Pound, J.D. (1998). IgG-Fc-mediated effector functions: molecular definition of interaction sites for effector ligands and the role of glycosylation. *Immunol. Rev.* 163, 59–76.

Jeong, H.-H., Kim, Y.-G., Jang, S.-C., Yi, H., and Lee, C.-S. (2012). Profiling surface glycans on live cells and tissues using quantum dot-lectin nanoconjugates. *Lab. Chip* 12, 3290–3295.

K

Kabsch, W. (2010). XDS. *Acta Crystallogr. D Biol. Crystallogr.* 66, 125–132.

Kadokura, H., Katzen, F., and Beckwith, J. (2003). Protein Disulfide Bond Formation in Prokaryotes. *Annu. Rev. Biochem.* 72, 111–135.

- Karsten, U., and Goletz, S. (2015). What controls the expression of the core-1 (Thomsen—Friedenreich) glycotope on tumor cells? *Biochem. Mosc.* *80*, 801–807.
- Kaur, M., Singh, K., Rup, P.J., Kamboj, S.S., and Singh, J. (2009). Anti-insect potential of lectins from *Arisaema* species towards *Bactrocera cucurbitae*. *J. Environ. Biol.* *30*, 1019–1023.
- Kilpatrick, D.C. (2002). Animal lectins: a historical introduction and overview. *Biochim. Biophys. Acta BBA - Gen. Subj.* *1572*, 187–197.
- Kim, Y.-S., Son, O.L., Lee, J.Y., Kim, S.H., Oh, S., Lee, Y.S., Kim, C.-H., Yoo, J.S., Lee, J.-H., Miyoshi, E., et al. (2008). Lectin precipitation using phytohemagglutinin-L4 coupled to avidin–agarose for serological biomarker discovery in colorectal cancer. *PROTEOMICS* *8*, 3229–3235.
- Kizuka, Y., Nakano, M., Yamaguchi, Y., Nakajima, K., Oka, R., Sato, K., Ren, C.-T., Hsu, T.-L., Wong, C.-H., and Taniguchi, N. (2017). An Alkynyl-Fucose Halts Hepatoma Cell Migration and Invasion by Inhibiting GDP-Fucose-Synthesizing Enzyme FX, TSTA3. *Cell Chem. Biol.* *24*, 1467–1478.e5.
- Kobayashi, Y., and Kawagishi, H. (2014). Fungal lectins: a growing family. *Methods Mol. Biol. Clifton NJ* *1200*, 15–38.
- Kobayashi, Y., Tateno, H., Dohra, H., Moriwaki, K., Miyoshi, E., Hirabayashi, J., and Kawagishi, H. (2012). A Novel Core Fucose-specific Lectin from the Mushroom *Pholiota squarrosa*. *J. Biol. Chem.* *287*, 33973–33982.
- Kornfeld, R., and Kornfeld, S. (1985). Assembly of asparagine-linked oligosaccharides. *Annu. Rev. Biochem.* *54*, 631–664.
- Kornfeld, S., Li, E., and Tabas, I. (1978). The synthesis of complex-type oligosaccharides. *J Biol Chem* *253*, 7771–7778.
- Kostrominova, T.Y. (2011). Application of WGA lectin staining for visualization of the connective tissue in skeletal muscle, bone and ligament/tendon studies. *Microsc. Res. Tech.* *74*, 18.
- Krissinel, E., and Henrick, K. (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D Biol. Crystallogr.* *60*, 2256–2268.
- Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* *372*, 774–797.
- Kusama, K., Okamoto, Y., Saito, K., Kasahara, T., Murata, T., Ueno, Y., Kobayashi, Y., Kamada, Y., and Miyoshi, E. (2017). Reevaluation of *Pholiota squarrosa* lectin-reactive haptoglobin as a pancreatic cancer biomarker using an improved ELISA system. *Glycoconj. J.* *34*, 537–544.
- Kuzmanov, U., Kosanam, H., and Diamandis, E.P. (2013). The sweet and sour of serological glycoprotein tumor biomarker quantification. *BMC Med.* *11*, 31.

L

- La Belle, J.T., Gerlach, J.Q., Svarovsky, S., and Joshi, L. (2007). Label-free impedimetric detection of glycan-lectin interactions. *Anal. Chem.* *79*, 6959–6964.

- Laemmli, U.K. (1970). Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* 227, 680.
- Lairson, L.L., Henrissat, B., Davies, G.J., and Withers, S.G. (2008). Glycosyltransferases: Structures, Functions, and Mechanisms. *Annu. Rev. Biochem.* 77, 521–555.
- Langer, G., Cohen, S.X., Lamzin, V.S., and Perrakis, A. (2008). Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat. Protoc.* 3, 1171–1179.
- Lee, Y.C., and Lee, R.T. (1995). Carbohydrate-Protein Interactions: Basis of Glycobiology. *Acc. Chem. Res.* 28, 321–327.
- Leonidas, D.D., Swamy, B.M., Hatzopoulos, G.N., Gonchigar, S.J., Chachadi, V.B., Inamdar, S.R., Zographos, S.E., and Oikonomakos, N.G. (2007). Structural Basis for the Carbohydrate Recognition of the *Sclerotium rolfsii* Lectin. *J. Mol. Biol.* 368, 1145–1161.
- Li, C., Simeone, D.M., Brenner, D.E., Anderson, M.A., Shedden, K.A., Ruffin, M.T., and Lubman, D.M. (2009). Pancreatic Cancer Serum Detection Using A Lectin/Glyco-Antibody Array Method. *J. Proteome Res.* 8, 483–492.
- Li, X., Guan, F., Li, D., Tan, Z., Yang, G., Wu, Y., and Huang, Z. (2016). Identification of aberrantly expressed glycans in gastric cancer by integrated lectin microarray and mass spectrometric analyses. *Oncotarget* 7, 87284.
- Li, Y., Tao, S.-C., Bova, G.S., Liu, A.Y., Chan, D.W., Zhu, H., and Zhang, H. (2011). Detection and Verification of Glycosylation Patterns of Glycoproteins from Clinical Specimens Using Lectin Microarrays and Lectin-based Immunosorbent Assays. *Anal. Chem.* 83, 8509–8516.
- Lin, Z., Simeone, D.M., Anderson, M.A., Brand, R.E., Xie, X., Shedden, K.A., Ruffin, M.T., and Lubman, D.M. (2011). A Mass Spectrometric Assay for Analysis of Haptoglobin Fucosylation in Pancreatic Cancer. *J. Proteome Res.* 10, 2602–2611.
- Lis, H., and Sharon, N. (1998). Lectins: carbohydrate-specific proteins that mediate cellular recognition. *Chem. Rev.* 98, 637–674.
- Liu, L. (2015). Antibody glycosylation and its impact on the pharmacokinetics and pharmacodynamics of monoclonal antibodies and Fc-fusion proteins. *J. Pharm. Sci.* 104, 1866–1884.
- Liu, Y., Xia, B., Gleason, T.J., Castañeda, U., He, M., Berry, G.T., and Fridovich-Keil, J.L. (2012). N- and O-linked glycosylation of total plasma glycoproteins in galactosemia. *Mol. Genet. Metab.* 106, 442–454.
- Locker, G.Y., Hamilton, S., Harris, J., Jessup, J.M., Kemeny, N., Macdonald, J.S., Somerfield, M.R., Hayes, D.F., Bast, R.C., and ASCO (2006). ASCO 2006 update of recommendations for the use of tumor markers in gastrointestinal cancer. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 24, 5313–5327.
- Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., and Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42, D490–D495.

Loris, R., Overberge, D. van, Dao-Thi, M.-H., Poortmans, F., Maene, N., and Wyns, L. (1994). Structural analysis of two crystal forms of lentil lectin at 1.8 Å resolution. *Proteins Struct. Funct. Bioinforma.* *20*, 330–346.

Luirink, J., and Sinning, I. (2004). SRP-mediated protein targeting: structure and function revisited. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* *1694*, 17–35.

Lv, X., Song, J., Xue, K., Li, Z., Li, M., Zahid, D., Cao, H., Wang, L., Song, W., Ma, T., et al. (2019). Core fucosylation of copper transporter 1 plays a crucial role in cisplatin-resistance of epithelial ovarian cancer by regulating drug uptake. *Mol. Carcinog.* *58*, 794–807.

M

Ma, T., Wang, Y., Jia, L., Shu, J., Yu, H., Du, H., Yang, J., Liang, Y., Chen, M., and Li, Z. (2019). Increased expression of core-fucosylated glycans in human lung squamous cell carcinoma. *RSC Adv.* *9*, 22064–22073.

Makyio, H., Shimabukuro, J., Suzuki, T., Imamura, A., Ishida, H., Kiso, M., Ando, H., and Kato, R. (2016). Six independent fucose-binding sites in the crystal structure of *Aspergillus oryzae* lectin. *Biochem. Biophys. Res. Commun.* *477*, 477–482.

de Marco, A. (2009). Strategies for successful recombinant expression of disulfide bond-dependent proteins in *Escherichia coli*. *Microb. Cell Factories* *8*, 26.

Matsumura, K., Higashida, K., Ishida, H., Hata, Y., Yamamoto, K., Shigeta, M., Mizuno-Horikawa, Y., Wang, X., Miyoshi, E., Gu, J., et al. (2007). Carbohydrate binding specificity of a fucose-specific lectin from *Aspergillus oryzae*: a novel probe for core fucose. *J. Biol. Chem.* *282*, 15700–15708.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). *Phaser* crystallographic software. *J. Appl. Crystallogr.* *40*, 658–674.

McCoy Jr, J.P., Varani, J., and Goldstein, I.J. (1983). Enzyme-linked lectin assay (ELLA): use of alkaline phosphatase-conjugated *Griffonia simplicifolia* B4 isolectin for the detection of α -D-galactopyranosyl end groups. *Anal. Biochem.* *130*, 437–444.

Michelakakis, H., Moraitou, M., Mavridou, I., and Dimitriou, E. (2009). Plasma lysosomal enzyme activities in congenital disorders of glycosylation, galactosemia and fructosemia. *Clin. Chim. Acta* *401*, 81–83.

Miyoshi, E., Moriwaki, K., and Nakagawa, T. (2008). Biological function of fucosylation in cancer biology. *J. Biochem. (Tokyo)* *143*, 725–729.

Molinari, M. (2007). N-glycan structure dictates extension of protein folding or onset of disposal. *Nat. Chem. Biol.* *3*, 313–320.

Monsigny, M., Mayer, R., and Roche, A.-C. (2000). Sugar-lectin interactions: sugar clusters, lectin multivalency and avidity. *Carbohydr. Lett.* *4*, 35–52.

Morgan, G.W., Kail, M., Hollinshead, M., and Vaux, D.J. (2013). Combined biochemical and cytological analysis of membrane trafficking using lectins. *Anal. Biochem.* *441*, 21–31.

Mori, K., Iida, S., Yamane-Ohnuki, N., Kanda, Y., Kuni-Kamochi, R., Nakano, R., Imai-Nishiya, H., Okazaki, A., Shinkawa, T., Natsume, A., et al. (2007). Non-fucosylated therapeutic antibodies: the next generation of therapeutic antibodies. *Cytotechnology* *55*, 109–114.

Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., and Vagin, A.A. (2011). *REFMAC 5* for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* *67*, 355–367.

N

Nagre, N.N., Chachadi, V.B., Eligar, S.M., Shubhada, C., Pujari, R., Shastry, P., Swamy, B.M., and Inamdar, S.R. (2010). Purification and Characterization of a Mitogenic Lectin from *Cephalosporium*, a Pathogenic Fungus Causing Mycotic Keratitis. *Biochem. Res. Int.* *2010*, 1–6.

Nakagawa, T., Miyoshi, E., Yakushijin, T., Hiramatsu, N., Igura, T., Hayashi, N., Taniguchi, N., and Kondo, A. (2008). Glycomic Analysis of Alpha-Fetoprotein L3 in Hepatoma Cell Lines and Hepatocellular Carcinoma Patients. *J. Proteome Res.* *7*, 2222–2233.

Nakamura-Tsuruta, S., Kominami, J., Kuno, A., and Hirabayashi, J. (2006). Evidence that *Agaricus bisporus* agglutinin (ABA) has dual sugar-binding specificity. *Biochem. Biophys. Res. Commun.* *347*, 215–220.

do Nascimento, A.S.F., Serna, S., Beloqui, A., Arda, A., Sampaio, A.H., Walcher, J., Ott, D., Unverzagt, C., Reichardt, N.-C., Jimenez-Barbero, J., et al. (2015). Algal lectin binding to core (α 1–6) fucosylated N-glycans: Structural basis for specificity and production of recombinant protein. *Glycobiology* *25*, 607–616.

O

Oda, Y., Senaha, T., Matsuno, Y., Nakajima, K., Naka, R., Kinoshita, M., Honda, E., Furuta, I., and Kakehi, K. (2003). A New Fungal Lectin Recognizing α (1–6)-linked Fucose in the N-Glycan. *J. Biol. Chem.* *278*, 32439–32447.

Ohtsubo, K., and Marth, J.D. (2006). Glycosylation in Cellular Mechanisms of Health and Disease. *Cell* *126*, 855–867.

Okada, M., Chikuma, S., Kondo, T., Hibino, S., Machiyama, H., Yokosuka, T., Nakano, M., and Yoshimura, A. (2017). Blockage of Core Fucosylation Reduces Cell-Surface Expression of PD-1 and Promotes Anti-tumor Immune Responses of T Cells. *Cell Rep.* *20*, 1017–1028.

Okuyama, M., Ueno, H., Kobayashi, Y., Kawagishi, H., Takahashi, D., and Toshima, K. (2016). Target-selective photo-degradation of AFP-L3 and selective photo-cytotoxicity against HuH-7 hepatocarcinoma cells using an anthraquinone–PhoSL hybrid. *Chem. Commun.* *52*, 2169–2172.

Oliveira, C., Teixeira, J.A., and Domingues, L. (2013). Recombinant lectins: an array of tailor-made glycan-interaction biosynthetic tools. *Crit. Rev. Biotechnol.* *33*, 66–80.

Opdenakker, G., Rudd, P.M., Ponting, C.P., and Dwek, R.A. (1993). Concepts and principles of glycobiology. *FASEB J.* *7*, 1330–1337.

P

Peppas, V.I., Venkat, H., Kantsadi, A.L., Inamdar, S.R., Bhat, G.G., Eligar, S., Shivanand, A., Chachadi, V.B., Satisha, G.J., Swamy, B.M., et al. (2015). Molecular Cloning, Carbohydrate Specificity and the Crystal Structure of Two *Sclerotium rolfisii* Lectin Variants. *Molecules* *20*, 10848–10865.

Perez, S. (2018). Symbolic Representation of Monosaccharides in the Age of Glycobiology. <http://www.glycopedia.eu>.

Pervin, M., Koyama, Y., Isemura, M., and Nakamura, Y. (2015). Plant lectins in therapeutic and diagnostic cancer research. *Int J Plant Biol Res* 3, 1030–1035.

Pilobello, K.T., Krishnamoorthy, L., Slawek, D., and Mahal, L.K. (2005). Development of a lectin microarray for the rapid analysis of protein glycopatterns. *Chembiochem Eur. J. Chem. Biol.* 6, 985–989.

Pinho, S.S., and Reis, C.A. (2015). Glycosylation in cancer: mechanisms and clinical implications. *Nat. Rev. Cancer* 15, 540–555.

Prasthofer, T., Phillips, S.R., Suddath, F.L., and Engler, J.A. (1989). Design, expression, and crystallization of recombinant lectin from the garden pea (*Pisum sativum*). *J. Biol. Chem.* 264, 6793–6796.

Preithner, S., Elm, S., Lippold, S., Locher, M., Wolf, A., Silva, A.J. da, Baeuerle, P.A., and Prang, N.S. (2006). High concentrations of therapeutic IgG1 antibodies are needed to compensate for inhibition of antibody-dependent cellular cytotoxicity by excess endogenous immunoglobulin G. *Mol. Immunol.* 43, 1183–1193.

Prestegard, J.H., Liu, J., and Widmalm, G. (2015). Oligosaccharides and Polysaccharides. In *Essentials of Glycobiology*, A. Varki, R.D. Cummings, J.D. Esko, P. Stanley, G.W. Hart, M. Aebi, A.G. Darvill, T. Kinoshita, N.H. Packer, J.H. Prestegard, et al., eds. (Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press), pp. 31–40.

Prinz, W.A., \AAslund, F., Holmgren, A., and Beckwith, J. (1997). The Role of the Thioredoxin and Glutaredoxin Pathways in Reducing Protein Disulfide Bonds in the *Escherichia coli* Cytoplasm. *J. Biol. Chem.* 272, 15661–15667.

Q

Qiu, Y., Patwa, T.H., Xu, L., Shedden, K., Misek, D.E., Tuck, M., Jin, G., Ruffin, M.T., Turgeon, D.K., Synal, S., et al. (2008). Plasma Glycoprotein Profiling for Colorectal Cancer Biomarker Identification by Lectin Glycoarray and Lectin Blot. *J. Proteome Res.* 7, 1693–1703.

R

Raemaekers, R.J.M., Muro, L. de, Gatehouse, J.A., and Fordham-Skelton, A.P. (1999). Functional phytohemagglutinin (PHA) and *Galanthus nivalis* agglutinin (GNA) expressed in *Pichia pastoris*. *Eur. J. Biochem.* 265, 394–403.

Reily, C., Stewart, T.J., Renfrow, M.B., and Novak, J. (2019). Glycosylation in health and disease. *Nat. Rev. Nephrol.* 15, 346–366.

Reis, C.A., Osorio, H., Silva, L., Gomes, C., and David, L. (2010). Alterations in glycosylation as biomarkers for cancer detection. *J. Clin. Pathol.* 63, 322–329.

Rheinnecker, M., Hardt, C., Ilag, L.L., Kufer, P., Gruber, R., Hoess, A., Lupas, A., Rottenberger, C., Plückthun, A., and Pack, P. (1996). Multivalent antibody fragments with high functional affinity for a tumor-associated carbohydrate antigen. *J. Immunol.* 157, 2989–2997.

Ritz, D., Lim, J., Reynolds, C.M., Poole, L.B., and Beckwith, J. (2001). Conversion of a Peroxiredoxin into a Disulfide Reductase by a Triplet Repeat Expansion. *Science* 294, 158–160.

Rivinoja, A., Kokkonen, N., Kellokumpu, I., and Kellokumpu, S. (2006). Elevated Golgi pH in breast and colorectal cancer cells correlates with the expression of oncofetal carbohydrate T-antigen. *J. Cell. Physiol.* 208, 167–174.

Robert, X., and Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* 42, W320–W324.

Ronin, C., Fenouillet, E., Hovsepian, S., Fayet, G., and Fournet, B. (1986). Regulation of thyroglobulin glycosylation. A comparative study of the thyroglobulins from porcine thyroid glands and follicles in serum-free culture. *J. Biol. Chem.* 261, 7287–7293.

Rüdiger, H., and Gabius, H.-J. (2001). Plant lectins: Occurrence, biochemistry, functions and applications. *Glycoconj. J.* 18, 589–613.

S

Saldova, R., Fan, Y., Fitzpatrick, J.M., Watson, R.W.G., and Rudd, P.M. (2011). Core fucosylation and α 2-3 sialylation in serum N-glycome is significantly increased in prostate cancer comparing to benign prostate hyperplasia. *Glycobiology* 21, 195–205.

Satoh, M., Iida, S., and Shitara, K. (2006). Non-fucosylated therapeutic antibodies as next-generation therapeutic antibodies. *Expert Opin. Biol. Ther.* 6, 1161–1173.

Satoomaa, T., Heiskanen, A., Leonardsson, I., Ångström, J., Olonen, A., Blomqvist, M., Salovuori, N., Haglund, C., Teneberg, S., Natunen, J., et al. (2009). Analysis of the Human Cancer Glycome Identifies a Novel Group of Tumor-Associated N-Acetylglucosamine Glycan Antigens. *Cancer Res.* 69, 5811–5819.

Saxena, A., and Wu, D. (2016). Advances in Therapeutic Fc Engineering - Modulation of IgG-Associated Effector Functions and Serum Half-life. *Front. Immunol.* 7, 580.

Schägger, H. (2006). Tricine–SDS–PAGE. *Nat. Protoc.* 1, 16–22.

Schägger, H., and von Jagow, G. (1987). Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* 166, 368–379.

Schlegel, S., Rujas, E., Ytterberg, A.J., Zubarev, R.A., Luirink, J., and de Gier, J.-W. (2013). Optimizing heterologous protein production in the periplasm of *E. coli* by regulating gene expression levels. *Microb. Cell Factories* 12, 24.

Schmaltz, R.M., Hanson, S.R., and Wong, C.-H. (2011). Enzymes in the Synthesis of Glycoconjugates. *Chem. Rev.* 111, 4259–4307.

Schneider, S., and Nägele, E. (2014). Glycan characterization of the NIST RM monoclonal antibody using a total analytical solution: From sample preparation to data analysis. *Agilent Technologies*.

Schultz, M.J., Swindall, A.F., Wright, J.W., Sztul, E.S., Landen, C.N., and Bellis, S.L. (2013). ST6Gal-I sialyltransferase confers cisplatin resistance in ovarian tumor cells. *J. Ovarian Res.* 6, 25.

Sharon, N., and Lis, H. (1972). Lectins: Cell-Agglutinating and Sugar-Specific Proteins. *Science* 177, 949–959.

Sharon, N., and Lis, H. (2004). History of lectins: from hemagglutinins to biological recognition molecules. *Glycobiology* 14, 53R–62R.

Sheldrick, G.M. (2010). Experimental phasing with *SHELXC / D / E*: combining chain tracing with density modification. *Acta Crystallogr. D Biol. Crystallogr.* 66, 479–485.

Shinohara, Y., Hasegawa, Y., Kaku, H., and Shibuya, N. (1997). Elucidation of the mechanism enhancing the avidity of lectin with oligosaccharides on the solid phase surface. *Glycobiology* 7, 1201–1208.

Singh, R.S., Bhari, R., and Kaur, H.P. (2010). Mushroom lectins: Current status and future perspectives. *Crit. Rev. Biotechnol.* 30, 99–126.

Smith, B.J. (1994). SDS Polyacrylamide Gel Electrophoresis of Proteins. In *Basic Protein and Peptide Protocols*, (New Jersey: Humana Press), pp. 23–34.

Springer, G.F. (1984). T and Tn, general carcinoma autoantigens. *Science* 224, 1198–1206.

Stanley, P., Taniguchi, N., and Aebi, M. (2015). N-Glycans. In *Essentials of Glycobiology*, A. Varki, R.D. Cummings, J.D. Esko, P. Stanley, G.W. Hart, M. Aebi, A.G. Darvill, T. Kinoshita, N.H. Packer, J.H. Prestegard, et al., eds. (Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press), p. Chapter 9.

Steen, P.V. den, Rudd, P.M., Dwek, R.A., and Opdenakker, G. (1998). Concepts and principles of O-linked glycosylation. *Crit. Rev. Biochem. Mol. Biol.* 33, 151–208.

Sumisa, F., Ichijo, N., Yamaguchi, H., Nakatsumi, H., Ando, A., Iijima, N., Oguri, S., Uehara, K., and Nagata, Y. (2004). Molecular properties of mycelial aggregate-specific lectin of *Pleurotus cornucopiae*. *J. Biosci. Bioeng.* 98, 257–262.

Swanson, M.D., Winter, H.C., Goldstein, I.J., and Markovitz, D.M. (2010). A Lectin Isolated from Bananas Is a Potent Inhibitor of HIV Replication. *J. Biol. Chem.* 285, 8646–8655.

T

Tai, T., Yamashita, K., Ogata-Arakawa, M., Koide, N., Muramatsu, T., Iwashita, S., Inoue, Y., and Kobata, A. (1975). Structural studies of two ovalbumin glycopeptides in relation to the endo-beta-N-acetylglucosaminidase specificity. *J. Biol. Chem.* 250, 8569–8575.

Tait, A.R., and Straus, S.K. (2011). Overexpression and purification of U24 from human herpesvirus type-6 in *E. coli*: unconventional use of oxidizing environments with a maltose binding protein-hexahistidine dual tag to enhance membrane protein yield. *Microb. Cell Factories* 10, 51.

Takeya, A., Hosomi, O., Nishijima, H., Ohe, Y., Sugahara, K., Sagi, M., Yamazaki, K., Hayakawa, H., Takeshita, H., Sasaki, C., et al. (2007). Presence of β -linked GalNAc residues on N-glycans of human thyroglobulin. *Life Sci.* 80, 538–545.

Tateno, H., Nakamura-Tsuruta, S., and Hirabayashi, J. (2009). Comparative analysis of core-fucose-binding lectins from *Lens culinaris* and *Pisum sativum* using frontal affinity chromatography. *Glycobiology* 19, 527–536.

Tateno, H., Toyota, M., Saito, S., Onuma, Y., Ito, Y., Hiemori, K., Fukumura, M., Matsushima, A., Nakanishi, M., Ohnuma, K., et al. (2011). Glycome Diagnosis of Human Induced Pluripotent Stem Cells Using Lectin Microarray. *J. Biol. Chem.* *286*, 20345–20353.

Taylor, M.E., and Drickamer, K. (2003). Introduction to glycobiology (Oxford: Oxford Univ. Press).

Taylor, M.E., Drickamer, K., Schnaar, R.L., Etzler, M.E., and Varki, A. (2015). Discovery and Classification of Glycan-Binding Proteins. In *Essentials of Glycobiology*, A. Varki, R.D. Cummings, J.D. Esko, P. Stanley, G.W. Hart, M. Aebi, A.G. Darvill, T. Kinoshita, N.H. Packer, J.H. Prestegard, et al., eds. (Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press), p. Chapter 28.

Tegel, H., Ottosson, J., and Hober, S. (2011). Enhancing the protein production levels in *Escherichia coli* with a strong promoter. *FEBS J.* *278*, 729–739.

Terpe, K. (2006). Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Appl. Microbiol. Biotechnol.* *72*, 211–222.

Tian, E., and Ten Hagen, K.G. (2009). Recent insights into the biological roles of mucin-type O-glycosylation. *Glycoconj. J.* *26*, 325–334.

Tobias, R., and Kumaraswamy, S. (2014). Biosensors: fundamentals and applications. *ForteBio Appl. Note* *14*, 1–21.

Trigueros, V., Lougarre, A., Djamila, A.A., Rahbe, Y., Guillot, J., Chavant, L., Fournier, D., and Paquereau, L. (2003). *Xerocomus chrysenteron* lectin: Identification of a new pesticidal protein. *Biochim. Biophys. Acta* *1621*, 292–298.

Tu, C.-F., Wu, M.-Y., Lin, Y.-C., Kannagi, R., and Yang, R.-B. (2017). FUT8 promotes breast cancer cell invasiveness by remodeling TGF- β receptor core fucosylation. *Breast Cancer Res.* *19*, 111.

U

Ünligil, U.M., and Rini, J.M. (2000). Glycosyltransferase structure and mechanism. *Curr. Opin. Struct. Biol.* *10*, 510–517.

V

Vainauskas, S., Duke, R.M., McFarland, J., McClung, C., Ruse, C., and Taron, C.H. (2016). Profiling of core fucosylated N-glycans using a novel bacterial lectin that specifically recognizes α 1,6 fucosylated chitobiose. *Sci. Rep.* *6*.

Van Schaftingen, E., and Jaeken, J. (1995). Phosphomannomutase deficiency is a cause of carbohydrate-deficient glycoprotein syndrome type I. *FEBS Lett.* *377*, 318–320.

Vanti, G.L., Vishwanathreddy, H., Venkat, H., Bhat, G.G., Padmanabhan, V., Jayaprakash, N.S., Hiremath, V., Katageri, I.S., Gonchigar, S.J., Inamdar, S.R., et al. (2016). *Sclerotium rolfsii* lectin expressed in tobacco confers protection against *Spodoptera litura* and *Myzus persicae*. *J. Pest Sci.* *89*, 591–602.

Varki, A. (1993). Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* *3*, 97–130.

Varki, A., Cummings, R.D., Aebi, M., Packer, N.H., Seeberger, P.H., Esko, J.D., Stanley, P., Hart, G., Darvill, A., Kinoshita, T., et al. (2015). Symbol Nomenclature for Graphical Representations of Glycans. *Glycobiology* 25, 1323–1324.

Varrot, A., Basheer, S.M., and Imberty, A. (2013). Fungal lectins: structure, function and potential applications. *Curr. Opin. Struct. Biol.* 23, 678–685.

W

Wang, X., Inoue, S., Gu, J., Miyoshi, E., Noda, K., Li, W., Mizuno-Horikawa, Y., Nakano, M., Asahi, M., and Takahashi, M. (2005). Dysregulation of TGF- β 1 receptor activation leads to abnormal lung development and emphysema-like phenotype in core fucose-deficient mice. *Proc. Natl. Acad. Sci.* 102, 15791–15796.

Wang, X., Gu, J., Ihara, H., Miyoshi, E., Honke, K., and Taniguchi, N. (2006). Core Fucosylation Regulates Epidermal Growth Factor Receptor-mediated Intracellular Signaling. *J. Biol. Chem.* 281, 2572–2577.

Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M., and Barton, G.J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinforma. Oxf. Engl.* 25, 1189–1191.

Wimmerova, M., Mitchell, E., Sanchez, J.-F., Gautier, C., and Imberty, A. (2003). Crystal structure of fungal lectin: six-bladed beta-propeller fold and novel fucose recognition mode for *Aleuria aurantia* lectin. *J. Biol. Chem.* 278, 27059–27067.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., et al. (2011). Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* 67, 235–242.

Wu, A.M., Wu, J.H., Tsai, M.-S., Hegde, G.V., Inamdar, S.R., Swamy, B.M., and Herp, A. (2001). Carbohydrate specificity of a lectin isolated from the fungus *Sclerotium rolfsii*. *Life Sci.* 69, 2039–2050.

Y

Yamane-Ohnuki, N., and Satoh, M. (2009). Production of therapeutic antibodies with controlled fucosylation. *MAbs* 1, 230–236.

Yamasaki, K., Yamasaki, T., and Tateno, H. (2018). The trimeric solution structure and fucose-binding mechanism of the core fucosylation-specific lectin PhoSL. *Sci. Rep.* 8, 7740.

Yamasaki, K., Kubota, T., Yamasaki, T., Nagashima, I., Shimizu, H., Terada, R.-I., Nishigami, H., Kang, J., Tateno, M., and Tateno, H. (2019). Structural basis for specific recognition of core fucosylation in N-glycans by *Pholiota squarrosa* lectin (PhoSL). *Glycobiology* 29, 576–587.

Yan, Q., Jiang, Z., Yang, S., Deng, W., and Han, L. (2005). A novel homodimeric lectin from *Astragalus mongholicus* with antifungal activity. *Arch. Biochem. Biophys.* 442, 72–81.

Yin, J., Li, G., Ren, X., and Herrler, G. (2007). Select what you need: A comparative evaluation of the advantages and limitations of frequently used expression systems for foreign genes. *J. Biotechnol.* 127, 335–347.

Yu, L., Fernig, D.G., Smith, J.A., Milton, J.D., and Rhodes, J.M. (1993). Reversible Inhibition of Proliferation of Epithelial Cell Lines by *Agaricus bisporus* (Edible Mushroom) Lectin. *Cancer Res.* 53, 4627–4632.

Yu, L.-G., Andrews, N., Zhao, Q., McKean, D., Williams, J.F., Connor, L.J., Gerasimenko, O.V., Hilkens, J., Hirabayashi, J., Kasai, K., et al. (2007). Galectin-3 interaction with Thomsen-Friedenreich disaccharide on cancer-associated MUC1 causes increased cancer cell endothelial adhesion. *J. Biol. Chem.* 282, 773–781.

Z

Zhao, Y., Itoh, S., Wang, X., Isaji, T., Miyoshi, E., Kariya, Y., Miyazaki, K., Kawasaki, N., Taniguchi, N., and Gu, J. (2006). Deletion of Core Fucosylation on $\alpha 3\beta 1$ Integrin Down-regulates Its Functions. *J. Biol. Chem.* 281, 38343–38350.

Zhao, Y.-P., Xu, X.-Y., Fang, M., Wang, H., You, Q., Yi, C.-H., Ji, J., Gu, X., Zhou, P.-T., Cheng, C., et al. (2014). Decreased Core-Fucosylation Contributes to Malignancy in Gastric Cancer. *PLOS ONE* 9, e94536.

VII.

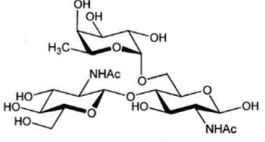
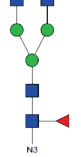
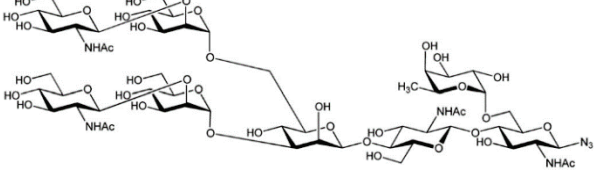
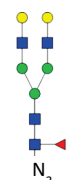
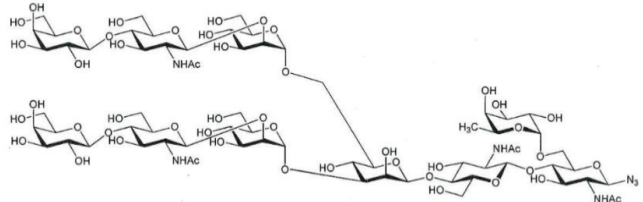
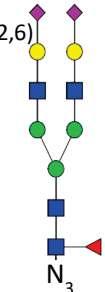
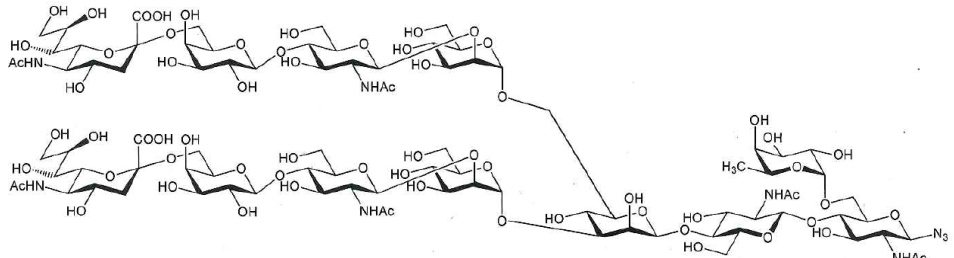
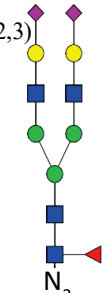
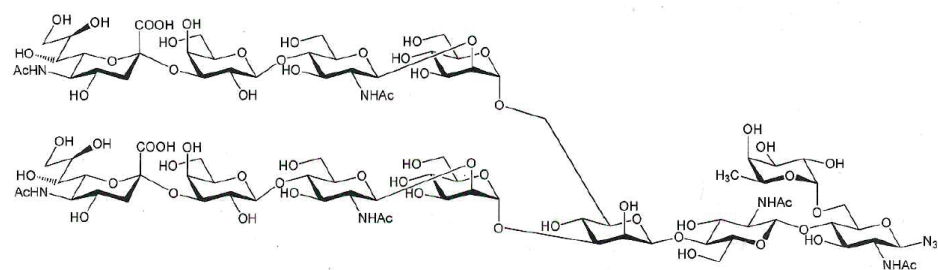
ANNEXES

1. Récapitulatif des séquences des oligonucléotides utilisés lors des expériences de mutagenèse dirigée

Amorce	Séquence
Mutagenèse de rPhoNL : délétion de la méthionine et de l'alanine en Nter	
rPhoNLm-For	5'ttcagggcgcgccggtgccggtgaccaaattgacctgcatggc3'
rPhoNLm-Rev	5'cgcgcgccctgaaaatacaggttttcggtaccagatctgggctgtcc3'
Mutagenèse de rPhoNLm : mutation de la sérine 70 en cystéine	
rPhoNLmSS-For	5'gacctgtgatggcgactgctatctgtcacagcg3'
rPhoNLmSS-Rev	5'cgctgtgcacagatagcagtcgccatcacaggtc3'
Mutagenèse de rPhoSL : délétion de la méthionine et de l'alanine en Nter	
rPhoSLm-For	5'ttcagggcgcgccggtgccggtgaccaaactggtgtgcatggc3'
rPhoSLm-Rev	5'cgcgcgccctgaaaatacaggttttcggtaccagatctgggctgtcc3'
Mutagenèse de rTrfbL2	
rTrfbL2-Y111A-For	5' gaaacagctctccaccgccactgtgaccagtgggc3'
rTrfbL2-Y111A-Rev	5'gccactggtcacagtggcggtggagagctgttc3'
rTrfbL2-R100G-For	5'gcagagtattacaattcagtcggctcgttcgcgcgagaaac3'
rTrfbL2-R100G-Rev	5'gttctcgcgcggaacgaaccgactgaattgtaatactctgc3'
rTrfbL2-Y111Q-For	5'gaaacagctctccaccagactgtgaccagtggg3'
rTrfbL2-Y111Q-Rev	5'cccactggtcacagtctgggtggagagctgttc3'
rTrfbL2-D77A-For	5'ctacaaacgctggggcgctattgttactggtctgtc3'
rTrfbL2-D77A-Rev	5'gacagaccagtaacaatagcggcccgctttgtag3'

Tableau VII-1. Récapitulatif des amorces utilisées au cours de l'étude. Les codons mutés sont indiqués en rouge.

2. Description détaillée des structures des N-glycanes utilisés dans cette étude

N-glycane	Structure
<p>Chitobiose Core- fucosylé</p> 	
<p>N-glycane 1</p> 	
<p>N-glycane 2</p> 	
<p>N-glycane 3</p> <p>(2,6)</p> 	
<p>N-glycane 4</p> <p>(2,3)</p> 	

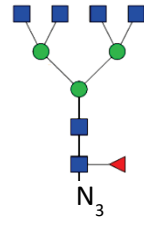
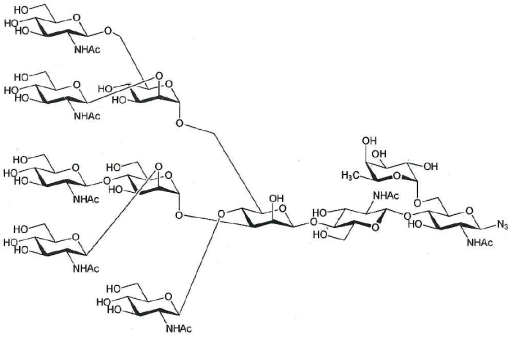
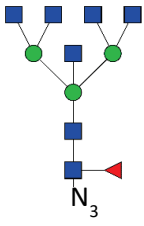
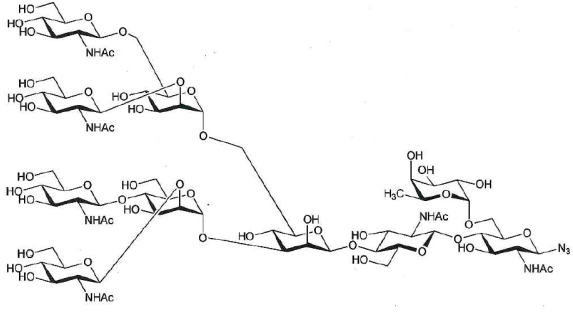
<p>N-glycane 5</p> 	
<p>N-glycane 6</p> 	

Tableau VII-2. Structures détaillées des N-glycanes utilisés dans l'étude

3. Représentation des N-glycanes 1, 2, 3, 5 et 6 en densité électronique

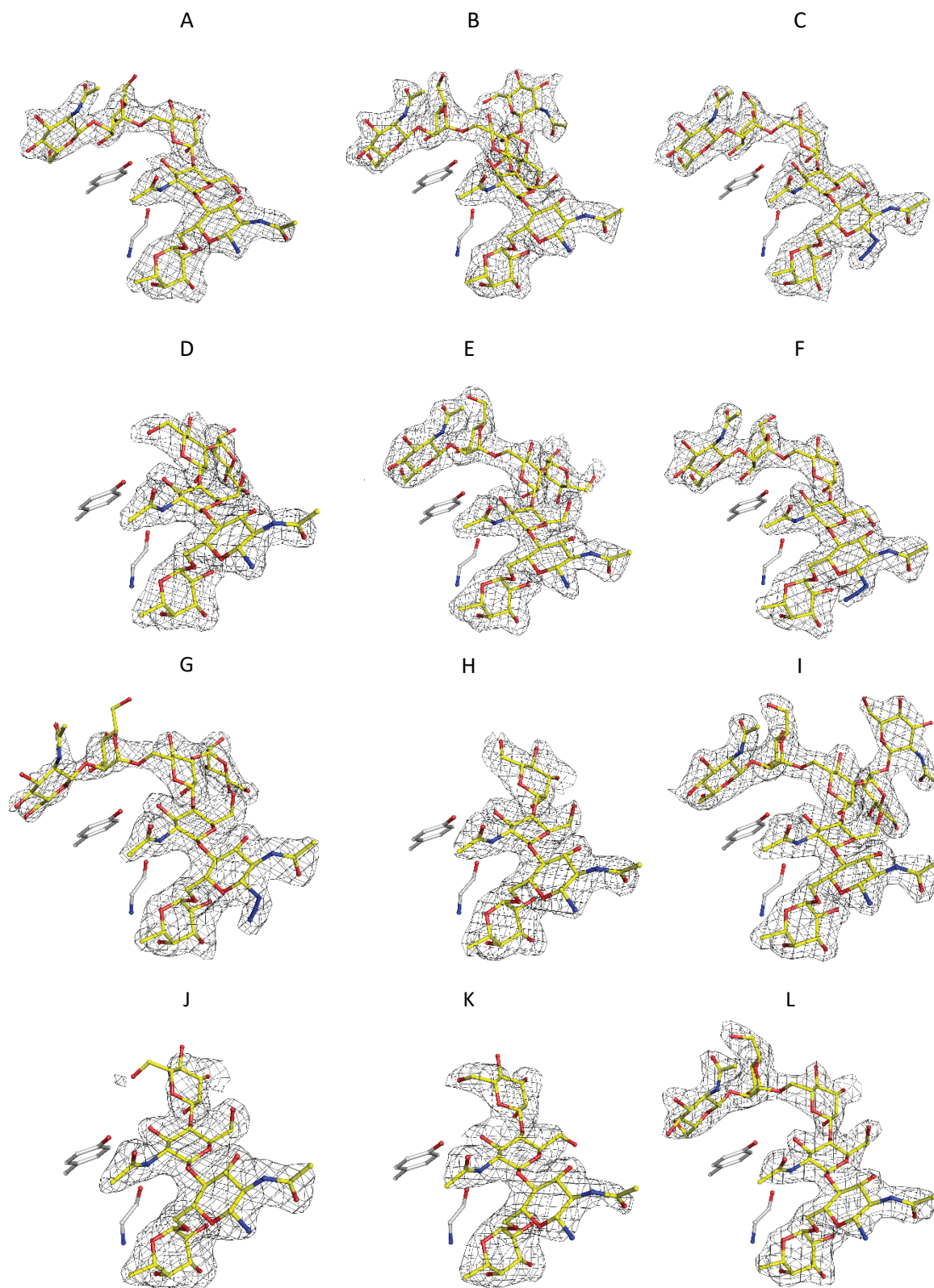


Figure VII-1. Densité électronique $2F_{\text{obs}}-F_{\text{calc}}$ du N-glycane 1 pour les 12 chaînes de rPhoSL affichée à 1σ .

La tyrosine 15 et la glycine 12 sont représentées et chaque figure porte le nom de la chaîne protéique du PDB 6FX1.

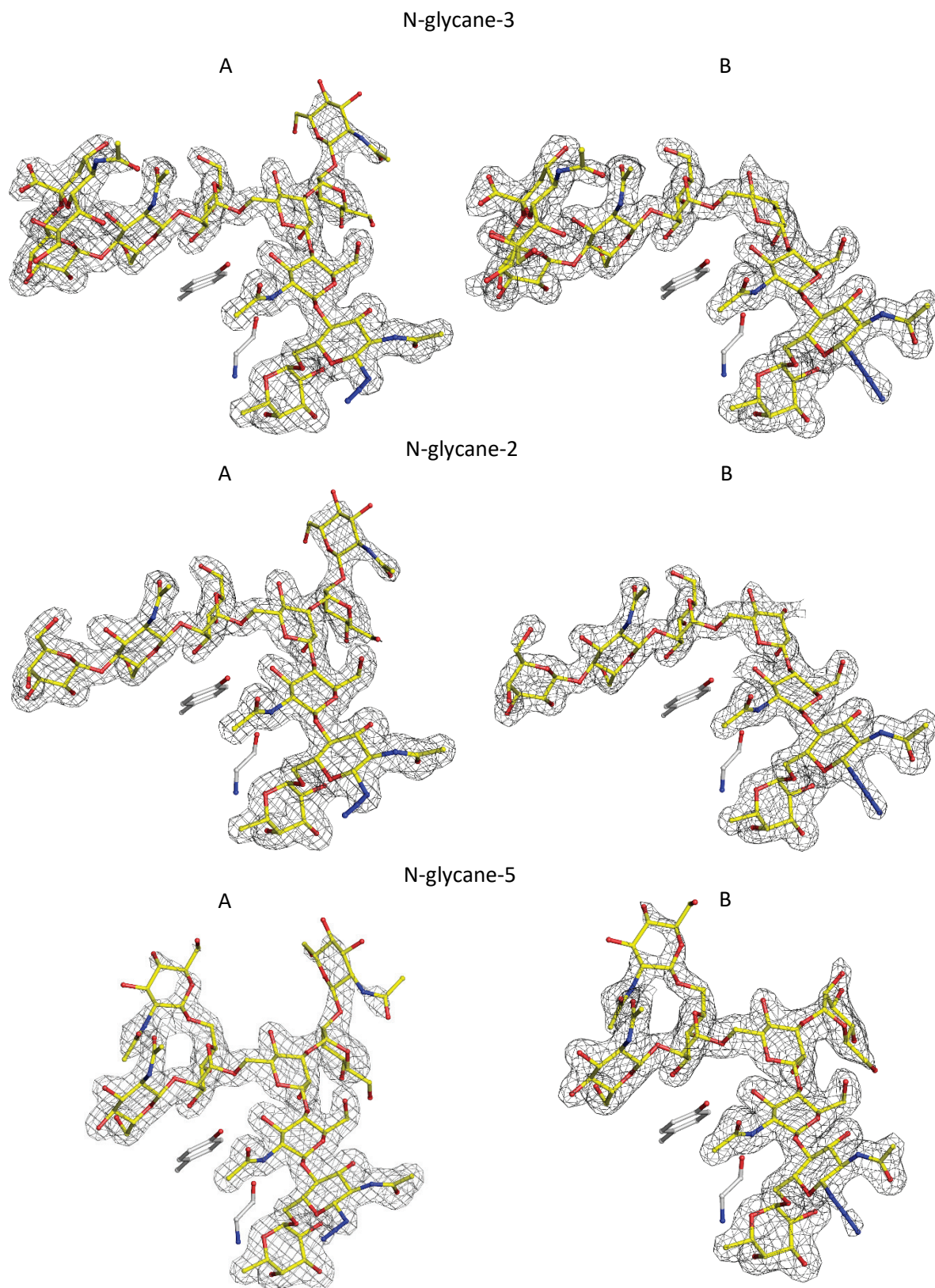


Figure VII-2. Densité électronique $2F_{\text{obs}}-F_{\text{calc}}$ des N-glycanes 2, 3 et 5 pour les chaînes A et B de rPhoSL affichée à 1σ .

La tyrosine 15 et la glycine 12 sont représentées et chaque figure porte le nom de la chaîne protéique du PDB correspondant.

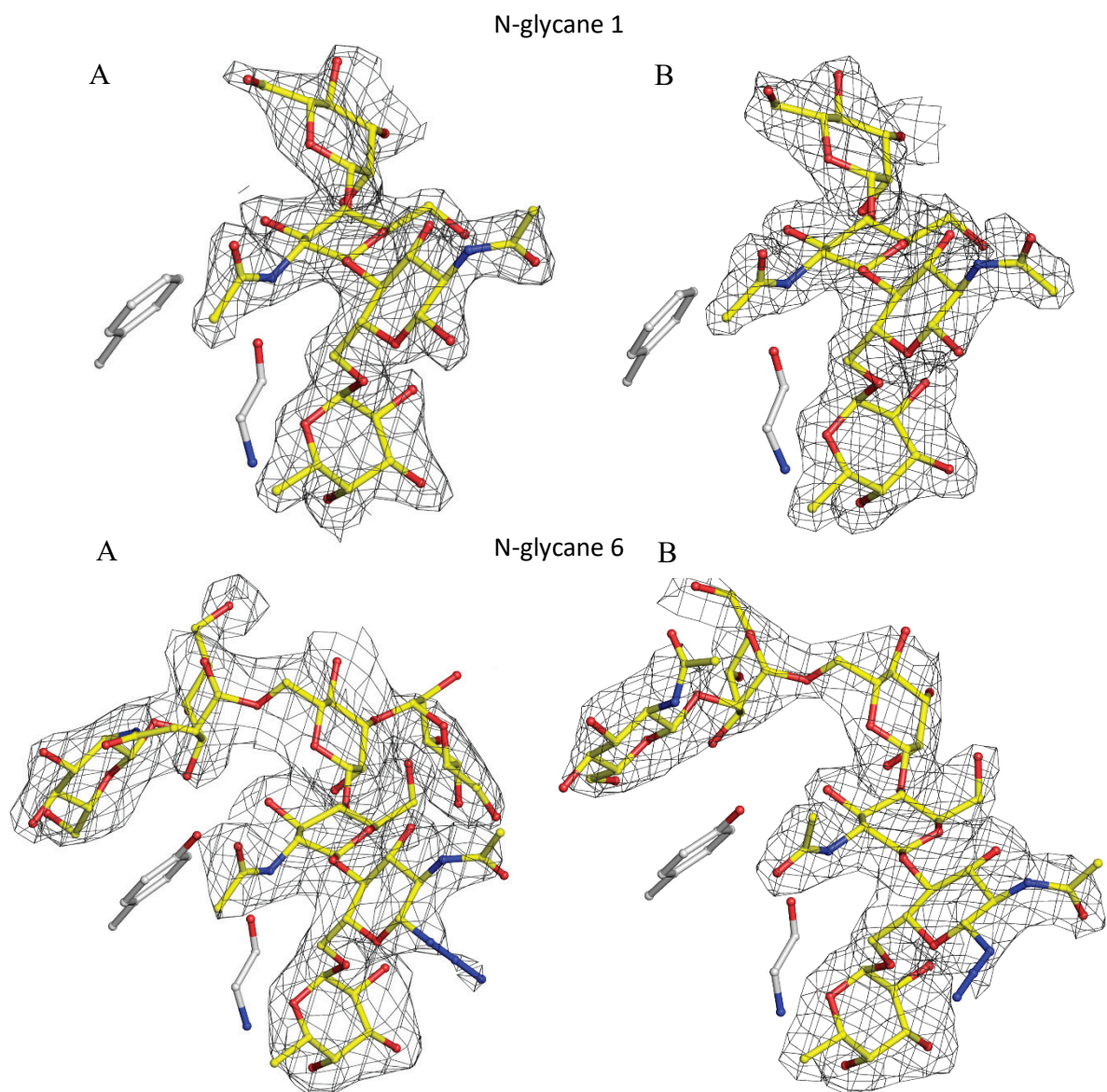


Figure VII-3. Densité électronique $2F_{obs}-F_{calc}$ du N-glycane 1 pour les chaînes A et B de rPhoNLSeMet et du N-glycane 6 pour les chaînes A et B de rPhonLm affichée à 1σ .
 La tyrosine ou phénylalanine 15 et la glycine 12 sont représentées et chaque figure porte le nom de la chaîne protéique du PDB correspondant.

4. Analyse des angles de torsion des N-glycanes

Complexe	N-glycane 1												N-glycane 2		N-glycane 3		N-glycane 5	
Chain	A	B	C	D	E	F	G	H	I	J	K	L	A	B	A	B	A	B
Fuc(α1,6)GlcNAc																		
Φ^*	-69	-66	-61	-62	-68	-64	-62	-67	-65	-59	-53	-60	-67	-58	-67	-65	-69	-54
Ψ^*	-142	-143	-145	-153	-140	-150	-141	-152	-145	-144	-151	-156	-145	-146	-144	-142	-143	-147
Ω^*	-59	-56	-55	-62	-65	-51	-61	-63	-66	-61	-70	-55	-60	-59	-59	-54	-60	-65
GlcNAc(β1,4)GlcNAc																		
Φ^*	-71	-66	-75	-75	-73	-74	-72	-79	-76	-71	-72	-80	-85	-81	-83	-78	-84	-82
Ψ^*	-132	-130	-131	-124	-136	-130	-136	-128	-134	-135	-128	-129	-141	-129	-143	-131	-143	-128
GlcNAc(β1,4)Man																		
Φ^*	-77	-73	-74	-76	-69	-78	-71	-78	-74	-74	-66	-71	-74	-86	-73	-77	-68	-82
Ψ^*	-141	-148	-130	-128	-125	-135	-127	-119	-124	-117	-126	-120	-139	-146	-141	-144	-137	-143
Man(α1,3)Man																		
Φ^*		65		83	73		54		-83				72		69		79	80
Ψ^*		132		131	103		143		131				133		129		138	137
Man(β1,2)GlcNAc sur la branche 1,3																		
Φ^*		-92							-78				-90		-89		-101	
Ψ^*		153							152				146		152		139	
Man(α1,6)Man																		
Φ^*	81	71	76		56	89	53		61			59	79	75	81	72	79	79

Ψ^*	-165	177	-164		172	-164	-178		168			171	-162	-162	-160	-159	-163	-167
Man(β1,2)GlcNAc sur la branche 1,6																		
Φ^*	-91	-84	-92		-94	-83	-72		-104			-105	-93	-87	-89	-84	-94	-89
Ψ^*	111	135	100		151	91	110		154			143	104	103	101	102	103	101
GlcNAc (β1,4)Gal sur la branche 1,6																		
Φ^*													-87	-87	-90	-90		
Ψ^*													-146	-148	-149	-148		
Sia(α2,6)Gal sur la branche 1,6																		
Φ^*															68	56		
Ψ^*															-179	-175		
GlcNAc(α1,6)Man																		
Φ^*																	-76	-71
Ψ^*																	-149	-149

Tableau VII-3. Tableau des angles de torsion pour les N-glycanes en interaction avec rPhoSL (1, 2, 3 et 5)

* ϕ ($O_5-C_1-O_x-C_x$) et ψ ($C_1-O_x-C_x-C_{x+1}$) avec x, le numéro de l'atome de carbone du second monosaccharide avec lequel la liaison glycosidique a lieu et Ω ($O_6-C_6-C_5-O_5$) pour deux monosaccharides avec une liaison 1 \rightarrow 6.

Article (Angewandte, 2018)

Recognition of Complex Core-Fucosylated N-glycans by a Mini Lectin

Aurore Cabanettes, Lukas Perkams, Carolina Spies, Prof. Dr. Carlo Unverzagt, Dr. Annabelle Varrot

Abstract

The mini fungal lectin PhoSL was recombinantly produced and characterized. Despite a length of only 40 amino acids, PhoSL exclusively recognizes N-glycans with α 1,6-linked fucose. Core fucosylation influences the intrinsic properties and bioactivities of mammalian N-glycoproteins and its level is linked to various cancers. Thus, PhoSL serves as a promising tool for glycoprofiling. Without structural precedence, the crystal structure was solved using the zinc anomalous signal, and revealed an interlaced trimer creating a novel protein fold termed β -prism III. Three biantennary core-fucosylated N-glycan azides of 8 to 12 sugars were cocrystallized with PhoSL. The resulting highly resolved structures gave a detailed view on how the exclusive recognition of α 1,6-fucosylated N-glycans by such a small protein occurs. This work also provided a protein consensus motif for the observed specificity as well as a glimpse into N-glycan flexibility upon binding.

<https://doi.org/10.1002/anie.201805165>

RÉSUMÉ

Les lectines sont des protéines ubiquitaires qui se lient spécifiquement et de manière réversible aux sucres sans les modifier. Elles peuvent déchiffrer le glycode : les informations stockées dans les sucres à la surface d'une cellule. Les lectines sont impliquées dans de nombreux processus biologiques tels que la cohésion cellulaire, le développement, la signalisation cellulaire, la défense ou les infections microbiennes. Elles peuvent présenter des propriétés exploitables, par exemple antiprolifératives, antivirales ou antitumorales et sont des outils recherchés en glycobiologie. Le règne fongique a suscité un large intérêt ces dernières années, car il est une source prometteuse et largement inexplorée pour les lectines ayant de nouvelles spécificités et/ou activités. Les données structurales sont encore limitées pour les lectines fongiques et ce projet vise notamment à combler cette lacune en étudiant deux familles de lectines reconnaissant des marqueurs glycosidiques associés aux cancers. Via une approche multidisciplinaire, l'interaction de ces lectines avec leur ligand au niveau atomique a été disséquée et des modifications ont été réalisées pour les rendre plus spécifiques pour un motif particulier. Les lectines recombinantes sauvages ou mutées ont été exprimées dans *Escherichia coli* puis purifiées par des méthodes chromatographiques classiques. La spécificité et l'affinité ont été déterminées par différentes techniques telles que l'hémagglutination, les puces à sucre et la microcalorimétrie de titration. La caractérisation structurale a été effectuée par cristallographie aux rayons X et a permis de définir les déterminants de la protéine et du ligand indispensables à l'interaction et d'aider à la conception et à l'évaluation structurale de mutants. Enfin, des analyses sur cellules et tissus cancéreux ont confirmé leur potentiel en tant qu'outils de marquage, ce qui permet d'envisager leur utilisation pour des applications biomédicales ou biotechnologiques.

Mots clés : lectines, biomarqueurs, champignon, cancer, biologie structurale

SUMMARY

Lectins are ubiquitous proteins which bind specifically and reversibly carbohydrates without modifying them. They can decipher the glycode: information stored in carbohydrates found at the surface of any cell. Lectins are involved in many biological processes such as cell cohesion, development, cell signaling, defense or microbial infections. They can present exploitable properties such as antiproliferative, antiviral or antitumor and are interesting tools in glycobiology. The fungal kingdom has attracted wide interest in the recent years since it is a promising and largely unexplored source for lectins for novel specificity and activities. Structural data are still limited for fungal lectins and this project aims to fill this gap of knowledge whilst studying new lectins specific for two glycosidic cancer markers. Via a multidisciplinary approach, we dissected the interaction of those lectins with their ligand at the atomic level or modified them to optimized their specificity or affinity. Recombinant wild-type and mutated lectins were expressed in *Escherichia coli* and then purified by classical chromatography methods. The specificity and affinity were determined by different techniques such as hemagglutination, glycan arrays, and microcalorimetry. Finally, structural characterization was performed by X-ray crystallography in order to localize the sugar binding site, to obtain the specific protein and ligand determinants necessary for the interaction and to help for the design and structural evaluation of mutants. Finally, analysis on cancer cells and tissues confirmed their potential as labelling tools thereafter for biomedical or biotechnological applications.

Keywords: lectins, biomarkers, mushroom, cancer, structural biology